



Pro gradu -tutkielma  
Fysiikan opettaja

# Dialogirakenne opettajaopiskelijoiden verkkokeskusteluissa kvanttifysiikasta

Miikka Turkkila  
8.8.2018

Ohjaajat: Prof. Ismo Koponen  
Dos. Maija Nousiainen

Tarkastajat: Prof. Ismo Koponen  
Dos. Maija Nousiainen

HELSINGIN YLIOPISTO  
FYSIKAN OSASTO

PL 64 (Gustaf Hällströmin katu 2a)  
00014 Helsingin yliopisto

Tiedekunta — Fakultet — Faculty		Osasto — Institution — Department	
Matemaattis-luonnontieteellinen tiedekunta		Fysiikan osasto	
Tekijä — Författare — Author			
Miikka Turkkila			
Työn nimi — Arbetets titel — Title			
Dialogirakenne opettajaopiskelijoiden verkkokeskusteluissa kvanttifysiikasta			
Oppiaine — Läroämne — Subject			
Fysiikan opettaja			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	
Pro gradu -tutkielma		8.8.2018	
		Sivumäärä — Sidoantal — Number of pages	
		68	
Tiivistelmä — Referat — Abstract			
<p>Tämän työn tarkoituksena oli kehittää työkalut verkkokeskustelun auki purkamiseen. Työkalut mahdollistaisivat verkkokeskustelun nopean analysoimisen, jolloin sitä voitaisiin käyttää mm. verkkooppimisen sosiaalisten ulottuvuuksien ymmärtämiseksi, kun dialogirakennetta verrataan teorioihin sosiaalisista rakenteista ja vuorovaikutuksista. Lisäksi työkaluja voidaan käyttää opetuksen tutkimuksen tukena ja verkkopohjaisen opetuksen kehittämisessä.</p> <p>Teoriataustana toimii tutkimus verkko-oppimisesta ja erityisenä aihealueena toimii tietokoneavusteinen kollaboratiivinen oppiminen. Termin alle jää laaja joukko eri opetustoimintaa, mutta tässä työssä termillä tarkoitetaan verkon välityksellä toteutettua ryhmäkeskustelua, jossa tavoitteena on oppiminen.</p> <p>Materiaalina toimi yhteensä 16 verkkokeskustelua, joissa neljä eri neljän hengen ryhmää keskustelivat neljästä eri kvanttifysiikkaan liittyvästä aiheesta. Tutkimusmenetelmänä käytettiin temaattista analyysiä keskustelun sisällön ja rakenteen varmistamiseksi. Tätä seurasi sosiaalinen verkostanalyysi keskustelun rakenteen kautta. Tähän käytettiin erityisesti McDonnell at. al (2014) kehittämää tapaa käyttää triadisia, eli kolmikoista muodostuneita, rooleja verkon analysoimiseksi.</p> <p>Analyysejä varten keskustelut taulukoitiin s.e. jokaisesta viestistä merkittiin lähettäjän ja lähetyksajan lisäksi kenelle viesti oli suunnattu ja mitä teemoja viesti piti sisällään. Tämän jälkeen kirjoitettiin Python-kieliset skriptit dialogirakenteen visualisoimiseksi ja niissä esiintyneiden roolien laskemiseksi.</p> <p>Tuloksiksi saatiin, että ryhmät keskustelivat tehtävänannon mukaisesti ja että verkkokeskustelun dialogirakenne voidaan esittää graafisesti niin sanottuna asynkronisena temporaalisena verkkona. Lisäksi keskusteluissa esiintyneet roolit voidaan laskea helposti ja esittää ns. lämpökartta-kuvana.</p> <p>Työn tavoitteet toteutuivat ja työssä kirjoitetut Python-skriptit lyhentävät merkittävästi verkko-keskustelun rakenteen analysoimista. Lisäksi tuloksia voidaan mahdollisesti käyttää ymmärtämään ryhmän sisäisiä sosiaalisia rakenteita. Tämä vaatii kuitenkin lisää työtä tässä käytetyn laskentamallin ja teorioiden yhdistämiseksi.</p>			
Avainsanat — Nyckelord — Keywords			
Verkkokeskustelu, verkko-oppiminen, sosiaalinen verkostanalyysi, temaattinen analyysi			
Säilytyspaikka — Förvaringsställe — Where deposited			
Helsingin yliopiston kirjasto, E-thesis			
Muita tietoja — Övriga uppgifter — Additional information			

# Sisältö

<b>1</b>	<b>Johdanto</b>	<b>1</b>
<b>2</b>	<b>Verkossa tapahtuva oppiminen</b>	<b>3</b>
2.1	Tietokoneavusteinen kollaboratiivinen oppiminen . . . . .	4
2.2	Asynkronisuus . . . . .	6
<b>3</b>	<b>Tutkimusmenetelmien teoriapohja</b>	<b>8</b>
3.1	Temaattinen analyysi . . . . .	8
3.2	Sosiaalinen verkostanalyysi . . . . .	10
3.2.1	Verkkoteorian perusteita . . . . .	10
3.2.2	Temporaaliset verkot . . . . .	11
3.2.3	Asynkroninen temporaalinen verkko . . . . .	12
3.3	Roolit . . . . .	13
3.3.1	Roolien yhteys sosiaalisiin verkkoihin . . . . .	17
<b>4</b>	<b>Tutkimus</b>	<b>20</b>
4.1	Tutkimuskysymykset . . . . .	20
4.2	Aineisto . . . . .	20
4.3	Toteutus . . . . .	22
4.3.1	Keskustelun rakenne ja temaattinen analyysi . . . . .	23
4.3.2	Python-skriptit dialogirakenteen ja roolisummien visualisointiin . . . . .	26
<b>5</b>	<b>Tulokset</b>	<b>29</b>
5.1	Teemat ja episodit . . . . .	29
5.2	Asynkroniset temporaaliset verkostot . . . . .	35
5.3	Roolit . . . . .	41
<b>6</b>	<b>Pohdinta</b>	<b>46</b>
	<b>Lähteet</b>	<b>50</b>
	<b>Liite A Python-skriptit</b>	<b>54</b>

# 1. Johdanto

Tämän työn taustalla on Helsingin yliopiston fysiikan osaston aineenopettajakoulutuksessa tehtävä opetuksen tutkimus ja kehittämistyö. Aiemmin on tutkittu fysiikan käsitteiden ja käsitteellisten rakenteiden oppimista. Tästä on muodostunut opetuksen jäsentävä lähestymistapa, jossa mm. didaktisilla rekonstruktioilla on suuri merkitys (Koponen, 2014). Jatkossa työ tulee myös osaksi keskittymään oppimisen sosiaalisiin ulottuvuuksiin.

Aineenopettajakoulutuksessa on huomattu, että tullessa fysiikan aineenopettaja opintoihin, opiskelijoiden minäpystyvyyden laskeminen. Taustalla tähän saattaa olla esimerkiksi se, että perus - ja aineopinnoissa opiskelu painottuu laskuharjoituksiin ja ilmiöiden matemaattiseen tarkasteluun. Opettajaopinnoissa taas tulisi ottaa haltuun fysiikan kokonaiskuva, ymmärtää ilmiöiden käsitteellisyttä tarkemmin ja paremmin, hahmottaa mitä milläkin käsitteillä todellisuudessa tarkoitetaan, miten ne eroavat toisistaan, miten käsitteet riippuvat toisistaan ja miten niistä rakentuu eheä käsittekokonaisuus.

Haasteellista tämä on mm. siksi, että fysiikassa yhdellä sanalla tai käsitteellä saataan tarkoittaa eri asioita. Esimerkiksi voima on sekä ilmiö, ominaisuus että suure. Toisaalta hiukkanen tarkoittaa eri asiaa klassisessa fysiikassa, kvanttimekaniikassa ja kvanttikenttäteoriassa. Asiaa ei helpota yhtään se, että mitä tarkalleen ottaen tarkoitetaan, pitää päätellä asiayhteydestä. Lisäksi useiden käsitteiden epämääräinen käyttö arkikielessä hämärtää niiden todellista merkitystä.

Se, että käsitteet eivät ole aina yksiselitteisiä, aiheuttaa myös ongelmia opetukseen ja oppimiseen. Oppimistapahtumassa syntyy helposti tilanteita, joissa puhutaan samasta asiasta tai ilmiöstä eri termeillä tai ymmärretään sama termi eri tavalla. Tämä saattaa aiheuttaa väärinymmärryksiä ja virhekäsityksiä etenkin jos opettajan tai luennoitsijan ja opiskelijoiden välillä on ero asian käsitteellisessä ymmärryksessä.

Etenkin kvantittumisen oppimiseen ja opettamiseen liittyy haasteita. Tässä yhteydessä puhutaan usein episteemisistä ja ontologisista ongelmista. Kvanttifysiikassa olioiden ja niistä saatavan tiedon luonne ei ole enää yhtä suoraviivainen ja selkeä kuin klassisessa fysiikassa. Tämä vaatiikin uudenlaista ajattelua ja abstraktin tieteellisen tiedon oppimista.

On tuotteliasta ja perusteltua saattaa opiskelijat keskustelemaan asioista keskenään. Tällöin he ikään kuin puhuvat fysiikkaa samalla kielellä. Lisäksi opiskelijoiden välisellä

vuoropuhelulla voi olla vaikutus vertaisarvostukseen ja minäpystyvyyteen (Koponen & Nousiainen, 2016), jotka edelleen vaikuttavat heidän väliseen viestintään (Bandura, 1997). Lisäksi internet-pohjaisessa oppimisessa minäpystyvyydellä on positiivinen vaikutus oppimistuloksiin (Tsai, Chuang, Liang, & Tsai, 2011).

Perinteisten ryhmätöiden ja kasvokkain keskustelun lisäksi keskustelua voidaan käyttää internetin välityksellä esimerkiksi chat-huoneissa tai keskustelupalstoilla. Tällöin puhutaan tietokoneavusteisesta kollaboratiivisesta oppimisesta, joka toimii tämänkin työn viitekehyksenä. Toisena teoreettisena lähtökohtana pidetään sosiaalisia verkkoja ja niiden rakennetta tutkivaa sosiaalista verkostanalyysiä.

Ei voida suoraan väittää, että verkkokeskustelussa esiintyvät rakenteet vastaavat ryhmän sosiaalisia rakenteita. Olemassa olevien sosiaalisten rakenteiden tiedetään kuitenkin vaikuttavan ihmisten välisiin vuorovaikutuksiin ja viestintään. Lisäksi vuorovaikutukset osaltaan muokkaavat sosiaalisia rakenteita ja verkostoja (E. J. Lawler, 2001; Skvoretz, Faust, & Fararo, 1996).

Sosiaalisten vuorovaikutusverkkojen ja rakenteiden syntymiselle ja kehittymiselle on olemassa erilaisia malleja. Onkin mielenkiintoista selvittää esiintyykö keskusteluissa niitä rakenteita, joita teorialla sosiaalisista rakenteista ennustavat. Jos samoja rakenteita löytyy tai ennemmin, jos on olemassa työkalu millä ne saadaan esille, mahdollistaisi se uusien tutkimuksien suunnittelun. Olisi esimerkiksi mahdollista verrata esiintyviä rakenteita oppimistuloksiin tai lähteä tarkastelemaan, onko vaikkapa opiskelijoiden välisellä vertaisarvostuksella vaikutusta dialogirakenteeseen.

Ennen kuin voidaan tarkastella keskusteluissa esiintyviä rakenteita, tulee varmistaa että keskustelu on ollut aiheenmukaista ja tehtäväorientoitunutta. Jos verkkokeskustelut eivät olisi tarpeeksi pitkiä eikä niissä keskusteltaisi aiheesta, ei keskustelun rakenteella olisi mielekästä merkitystä eikä sitä myöten yhteyttä työn muihin tavoitteisiin. Lisäksi tarkastelemalla viestien sisältöä varmistetaan, että foorumilla esiintyvä rakenne vastaa todellisia vuorovaikutuksia.

Työssä pyritään löytämään työkalut opiskelijoiden välisen verkkokeskustelun analysoimiseen tarkastelemalla opiskelijoiden välisiä vuorovaikutussuhteita dialogirakenteen mallintamiseksi. Tämän avulla on mahdollista saada parempi kuva oppimisen sosiaalisista ulottuvuuksista, joiden avulla opettamista voidaan kehittää edelleen. Tuotettujen työkalujen avulla on toivottavasti myös mahdollista arvioida verkkokeskusteluiden opetuksellista laatua ja esimerkiksi keskusteluiden ohjeistusta eri kurssien välillä.

## 2. Verkossa tapahtuva oppiminen

Internetin kehityksen myötä myös etäopetus on siirtynyt verkkoon (Anderson & Simpson, 2012). Verkossa tapahtuva oppiminen kattaa alleen laajan joukon eri opetus- ja oppimistapoja. Se voi tarkoittaa täysin etänä internetin välityksellä tapahtuvia verkkokursseja tai sekakursseja, joissa on sekä kontakti- että etäopetusta.

Verkkokurssi voi muistuttaa kirjekurssia, jossa oppimateriaalit ja tuotokset välitetään sähköisesti postin sijaan. Nykyään verkkokurssien kirjo on kuitenkin laaja ja ne mahdollistavat useat erilaiset aktiviteetit kuten vertaisviestinnän, videoluennot ja webkonferenssit, kohdennetut tehtävät ja tarkoitukseen tehdyt sovellusalueet sekä pelit. Verkkohaku on kuitenkin yksi useimmiten käytetyistä aktiviteeteistä verkko-oppimisessa (Tsai et al., 2011).

Internet ja sen mukana kehittyneet viestintäteknologiat mahdollistavat aktiivisen vuorovaikutuksen ihmisten välillä. Tämä on myös mahdollistanut verkkokursseille osallistuneiden opiskelijoiden keskinäisen vuorovaikutuksen passiivisen opiskelun sijaan.

Verkossa tapahtuvan vuorovaikutuksen tutkiminen alkoi 1970-luvulla tietokonevälitteisen viestinnän (computer-mediated communication, CMC) tutkimuksella (Laaksonen & Matikainen, 2013). Kyseinen termi on hyvin laaja käsite ja kattaakin lähes kaiken teknologian avulla tapahtuvan ihmisten välisen viestinnän. Käytännössä tutkimus on keskittynyt tekstipohjaisen aineiston analysointiin vaikka CMC-termin alla tapahtuva tutkimus on laajentunut teknologisen kehityksen myötä. Tämänlaisena teknologisenä kehityksenä voidaan pitää esimerkiksi muutosta anonyymeistä viestintäpalveluista sosiaaliseen mediaan.

CMC ei siis liity pelkästään oppimistarkoituksessa tapahtuvaan vuorovaikutukseen vaan kaikenlaiseen ihmistenväliseen vuorovaikutukseen. Se ei myöskään ota kantaa, onko viestintä kahden ihmisen vai suuremman ryhmän välistä. Termin sisältämä viestintä voi olla vastavuoroista tai esimerkiksi yhdensuuntaista massaviestintää.

Verkossa tapahtuva keskustelu voidaan jakaa kahteen eri päätyyppiin: synkroniseen ja asynkroniseen (Laaksonen & Matikainen, 2013). Synkronisella viestinnällä viitataan perinteisiin chat-keskusteluihin, joissa viestintä on reaaliaikaista. Esimerkiksi IRC-pikaviestintäpalvelu (Internet Relay Chat eli ”irkki”) on synkronisen verkkoviestinnän työkalu. Sähköposti edustaa taas asynkronista viestintää, jossa viestintä tapahtuu viiveellä.

Moderneissa verkkokeskusteluympäristöissä raja synkronisen ja asynkronisen välillä on hämärtynyt, kun viestintä voi olla hyvinkin nopea tempoista tai tapahtua selkeällä viiveellä. Rajaa hämärtää entisestään nykyisin käytössä olevat mobiilipalvelut, jotka mahdollistavat jatkuvan yhteyden verkkoon.

## 2.1 Tietokoneavusteinen kollaboratiivinen oppiminen

Kun tietokoneita käytetään yhteistyöhön oppimistarkoituksessa, puhutaan tietokoneavusteisesta kollaboratiivista oppimista (Computer-Suported Collaborative Learning, CSCL). Termille ei kuitenkaan ole yhtä selkeästi määriteltä tutkimusaluetta eikä asiaa helpota, että teknologian muuttuessa myös määritelmä muuttuu ja voi jopa olla ristiriidassa itsensä kanssa (Dillenbourg, Eurelings, & Hakkarainen, 2001). Termin alle jää niin erikseen tarkoitukseen rakennetut kuin vakiintuneetkin järjestelmät opetuksen käytössä. Toisin sanoen, siitä että tutkitaan oppimista CSCL kontekstissa, ei voida suoraan päätellä minkälaisessa ympäristössä tai millä työkaluilla opetus on järjestetty vaan se tulee aina varmistaa tapauskohtaisesti.

Käsite ei suoraan kerro minkälaisesta oppimistilanteesta on kyse, mutta kontekstista riippuen se voi olla luokassa tapahtuvaa (Iiskala, Volet, Lehtinen, & Vauras, 2015; Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003) tai yhteistyö on voitu toteuttaa etänä (Schellens & Valcke, 2005; Järvelä, Malmberg, & Koivuniemi, 2016; Schellens & Valcke, 2006). Vaikuttaisi kuitenkin siltä, että peruskoulun kontekstissa kyse on pääasiassa luokassa tapahtuvassa oppimisesta ja korkeakoulussa CSCL:llä tarkoitetaan verkkoviestintää luentojen ulkopuolella. Lisäksi yhteistyöhön on voitu käyttää sitä varten tuotettua ohjelmistoa, kuten CSILE (Scardamalia, Bereiter, Mclean, Swallow, & Woodruff, 1989) tai esimerkiksi korkeakoulun oman verkkoalustan foorumisovellusta, kuten esimerkiksi Helsingin yliopiston käyttämä Moodle-ympäristön keskusteluaktiiviteetti. Tietokonetta on siis käytetty joko oppimistyökaluna tai viestimiseen yhteistyön mahdollistamiseksi. Tämän perusteella CSCL ei välttämättä pidä sisällään tietokonevälitteistä viestintää, mutta usein ko. mahdollisuus on ainakin olemassa. Lisäksi joissain artikkeleissa on käytetty termiä CMC liitettyä oppimiseen yhteistyön kautta ilman erillistä mainintaa tai liittymistä tietokoneavusteiseen kollaboratiiviseen oppimiseen (Parke, 2017).

Tässä työssä CSCL:llä tarkoitetaan nimenomaan luokkahuoneen ulkopuolella tapahtuvaa yhteistyötä verkkokeskustelun avulla, jonka tavoitteena on oppiminen. Verkkokeskusteluissa oleellista on, että ajatukset tuodaan esille kirjoitetun tekstin muodossa. Tekstin tuottaminen vaatii opiskelijalta asian jäsentelyä ja selkeyttämistä, mutta toisaalta se mahdollistaa pidemmät puheenvuorot kuin mitkä olisi mahdollisia kasvokkain keskustel-

lessa (Pena-Shaffa & Nicholls, 2004). Toisaalta sähköisissä vuorovaikutuksissa ei voida välittää non-verbaaleja sosiaalisia merkkejä tai merkityksiä, mutta sähköisesti välitetään kuitenkin sellaista informaatiota, jota ei muuten välitettäisi (Sproull & Kiesler, 1986).

Kirjoitettu ja taltioitu teksti myös mahdollistaa opiskelijoiden reflektion jälkikäteen sekä palaamaan tarvittaessa keskusteluun muiden kurssitöiden apuna. Lisäksi hyvin organisoitu ja selkeästi ohjeistettu keskustelu mahdollistaa tekstien käytön tutkimuksen aineistona ja antaa siten opettajalle tavan kurssin kehittämiseen.

Alan tutkimuksessa analysoidaan opiskelijoiden tuottamia tekstejä, niiden osia tai niiden välistä yhteyttä (De Wever, Schellens, Valcke, & Van Keer, 2006). Aluksi tutkimus keskittyi keskustelun kvantitatiiviseen tarkasteluun mm. viestien määrästä, mutta nykyään aiheen tutkimuksessa on käytetään kvalitatiivisia ja kvantitatiivisia metodeja kollaboratiivisen oppimisen mallintamiseen ja analysoimiseen. Metodeja ovat mm. sisällönanalyysi, sosiaalinen verkostanalyysi, log-tiedostojen analysointi ja datan visuaaliset esitykset. Näitä on käytetty sekä oppimisprosessin, että -tulosten analysoimisessa (Puntambekar, Erkens, & Hmelo-Seilver, 2011). Lisäksi on voitu mitata yksilön tai koko ryhmän oppimista kokonaisuutena.

De Wever (2006) on tarkastellut CSCL:ssä käytettyjä sisällönanalyysimenetelmiä. Menetelmät ovat eronneet teoriataustan ja analyysiyksiköiden suhteen. Taustateorioina on ollut mm. sosiokonstruktivismi, kognitioteoria tai ”community of inquiry”. Analyysiyksikköinä on ollut yksittäinen lause tai kappale, viesti, temaattinen yksikkö tai koko keskustelu.

Edellä oleva osoittaa, että alan tutkimuksen käytänteissä on ollut eroja. Tämä ilmenee myös siinä, että vuorovaikutuksella on tarkoitettu eri asioita. Gunawardena et al (1997) onkin kehottanut vuorovaikutuksen määrittelemistä tarkasti. He itse määrittivät vuorovaikutuksen tiedon ”yhteisrakentamiseksi” viisivaiheisena prosessina kehittäessään Interaction Analysis Model (IAM) -työkalua. IAM on edelleen usein käytetty ja paljon viitattu sisällönanalyysimalli verkkokeskustelun analysoimiseksi.

Vuorovaikutus on myös ollut tekstien lukemista ja kirjoittamista (Daradoumis, Martinez-Mones, & Xhafa, 2006). Lisäksi esimerkiksi Discussion Analysis Tool (DAT) -työkalua on käytetty tunnistamaan vuorovaikutus rakenteita eli minkälainen vuorovaikutus seurasi mitäkin, eikä siis ihmisten välistä vuorovaikutuksia, ja mitkä niistä lisäsivät kriittistä ajattelua (Jeong, 2003).

Sisällönanalyysin ohella, tai lisäksi, usein käytetty menetelmä vuorovaikutuksen mallintamiseen on sosiaalinen verkostanalyysi. Sitä on käytetty esimerkiksi Gunawardenan mallin ohella ryhmän koheesion eli ryhmäläisten läsnäolon huomioimiseksi (Heo, Lim, & Kim, 2010) ja Daramoudis et al (2006) käyttivät SNA:ta ryhmän jäsenten aktiivisuuden mittaamiseen.

Usein tutkimuksen tavoitteina on ollut verkkoyhteistyön ymmärtäminen, kehittä-



minen tai hyötyjen osoittaminen. Yhteistyön uskotaan usein olevan hyödyllistä kaikille ryhmän jäsenille vaikka näin ei aina olekaan (Larson, 2010). Etenkin luovassa ideoinnissa yhteistyö kasvotusten saattaa heikentää prosessia passivoimalla osallistujia. Tietokonevälitteisessä kommunikaatiossa tämä ongelma lievenee, kun ryhmän jäsenet voivat muotoilla omia ideoitaan toisten nähtäväksi odottamatta omaa puheenvuoroaan. Hyvin järjestetty verkkoyhteistyö voikin mahdollisesti lisätä osallistujien aktiivisuutta, joka edelleen voi parantaa oppimistuloksia. Väitettä tukee havainnot siitä, että aktiiviset ryhmä menestyvät paremmin (Schellens & Valcke, 2005; Järvelä et al., 2016). Toisaalta viestien määrä ei itsessään takaa laadukasta verkkokeskustelua (Heo et al., 2010).

Onnistunut CSCL vaatii oikean ryhmäkoon, sillä pienet ryhmät heijastelevat korkeampia tiedonrakentumisen tasoja (Schellens & Valcke, 2006) ja isoimmissa ryhmissä muodostuu helposti isolaatioita eli joku ryhmä jäsenistä jää keskustelun ulkopuolelle (Koponen & Nousiainen, 2018). Toisaalta pienet ryhmät eivät hyödy verkkokeskustelusta, jos he tapaavat muutenkin, sillä tällöin vain vähän merkityksellistä tietoa julkaistaan verkkokeskustelussa (Wang, Woo, & Zhao, 2009). Tämä johtunee siitä, että jos ryhmän jäsenet tapaavat usein esimerkiksi muilla kursseilla, heidän on ainakin jossain määrin turha keskustella verkossa. Tällöin heidät on myös vaikea saada keskustelemaan verkossa ja vaarana on, että viestien lähettäminen nähdään vain pakollisena kurssisuorituksena.

Tutkimustulokset osoittavat myös, että verkkokeskustelut ovat tehtäväorientoitunutta, pysyy sellaisena ja reflektoi korkeita tiedonrakentumisen vaiheita (Schellens & Valcke, 2005). Kompleksinen ajattelu ja korkeaa tiedon rakentuminen vaikuttaisi olevan mahdollista (Lucas, Gunawardena, & Moreira, 2014) ja vuorovaikutukset jotka liittyivät eriäviin mielipiteisiin lisäsivät enemmän keskustelua ja kriittistä ajattelua (Jeong, 2003). Lisäksi Järvelä et. al (2006) ovat osoittaneet, että suunnitellut yhteistyön säätelyaktiviteetit jaetaan käytännön työskentelyssä ja ryhmät, jotka saavuttivat hyviä tuloksia käyttivät useita eri säätelyprosesseja oppimisen tukena sekä jakoivat säätelyä tehtävän aikana.

## 2.2 Asynkronisuus

Aivan kuten verkkoviestinnässä yleisesti, myös CSCL voidaan jakaa synkroniseen ja asynkroniseen. Edellä esitetyt hyödyt vaikuttaisivat koskevan etenkin juuri asynkronista keskustelua ja on osoitettu, että ns. asynkroninen oppimisverkosto on vähintään yhtä tehokas oppimismuoto kuin perinteinen luokkaopetus (Fjermestad & Hiltz, 2005). Myös monet tutkimukset käsittelevät nimen omaan asynkronista tietokoneavusteista kollaboratiivista oppimista.

Aiemmin todettiin, että raja synkronisen ja asynkronisen viestinnän välillä hämärtyy. Olisikin syytä tarkemmin määritellä asynkronisuus CSCL konseptissa. Yksi selkeä vaatimus asynkronisuudelle voisi olla selkeä mahdollisuus suoraan viitata edelliseen vies-

tiin, joka on mahdollista nykyaikaisissa verkkokeskusteluissa. Tämä generoikin niin kutsutun puumallin keskustelulle, jossa viestit esitetään temporaalisen rakenteen lisäksi hierarkisesti. Tällöin myöskään viestien välinen aika ei rajoita keskustelun määrittelyä.

Hiltz ja Goldman (2005) ovat määritelleet niin sanotun asynkronisen oppimisverkon (Asynchronous Learning Network, ALN) ihmisten väliseksi verkostoksi, jossa oppiminen voi tapahtua missä vain milloin vain. Juuri missä ja milloin vain tarkoittaa asynkronisuutta. Hiltzin ja Goldmanin sisältävät kirjassaan ALN termin alle myös kurssit, joissa on sekoitettuna verkko-oppimista ja kontaktiopetusta.

Tämän pohjalta määritellään asynkroninen CSCL sellaiseksi internetin välityksellä tapahtuvaksi keskusteluksi formaalissa oppimistarkoituksessa, joka ei riipu osallistujien fyysisestä sijainnista tai vastausajasta, mutta suurin osa viesteistä viittaa suoraan (teknisesti) aiempaan viestiin ja viestit jäävät kaikkien osallistujien nähtäväksi pysyvästi.

Asynkronisen verkko-oppimisen kaikista tärkein ominaisuus onkin, että jokainen henkilö voi työskennellä hänelle sopivaan tahtiin ja parhaiten sopivalla ajanhetkellä. Esimerkiksi ALN-tyyppisillä kursseilla kurssin jäsenet eivät tyypillisesti ole läsnä samaan aikaan samassa paikassa ja kurssit ovat jatkuvia kellon ympäri, jolloin opiskelijat lähettävät ja vastaanottavat viestejä heille sopivalla ajanhetkellä. Tyypillisesti opiskelija ei saa heti vastausta viestiinsä, mutta viimeistään seuraavalla sisäänkirjautumisella on nähtävillä toisen opiskelijan vastaus. Tämä saattaa aluksi vaikuttaa haitalta, mutta todellisuudessa siitä on vain hyötyä, sillä tällä tavoin ei muodostu aikatauluongelmia ihmisten välillä. Rauhallisempi tahti mahdollistaa myös syvällisemmän asian käsittelyn ja oppimisen (Hiltz & Goldman, 2005).

## 3. Tutkimusmenetelmien teoriapohja

Työssä tutkitaan opiskelijoiden välisiä vuorovaikutussuhteita dialogirakenteen mallintamiseksi. Tähän tarvitaan graafiteoriasta syntynyttä verkostanalyysia. Dialogirakenteita haluttiin tarkastella keskusteluepisodeittain, joten ennen verkostanalyysiä tehdään temaattinen analyysi. Temaattisen analyysin avulla voidaan tarkastella viestien sisältöjä, minkä perusteella keskustelut jaetaan episodeiksi.

### 3.1 Temaattinen analyysi

Psykologit Braun ja Clarke (2006) ovat määritelleet temaattisen analyysin laadullisen tutkimuksen metodiksi, jolla identifioidaan, analysoidaan ja raportoidaan rakenteellisia merkityksiä tai teemoja aineistosta. Analyysin avulla voidaan siis kuvailla ja organisoida yksityiskohtaista aineistoa.

Temaattisessa analyysissä aineistoon tutustaan ja siitä pyritään löytämään aineistoa kuvaavia koodeja. Tämän jälkeen aineistoista ja koodeista etsitään yhtenäisiä teemoja, joita myös tarkistetaan iteratiivisesti. Teemat määritellään ja nimetään lopullisesti, jonka jälkeen löytyneet teemat voidaan raportoida. Vaiheet ovat kuvattu tarkemmin taulukossa 3.1.

**Taulukko 3.1:** Temaattisen analyysin vaiheet (Braun & Clarke, 2006).

Vaihe	Vaiheen kuvaus
1. Aineistoon tutustuminen	Aineiston litterointi, lukeminen ja uudelleen lukeminen, alustava ideointi.
2. Alustava koodaus	Mielenkiintoisten koodien systemaattinen kokoaminen koko aineistosta.
3. Teemojen etsiminen (koodista)	Koodien järjestäminen potentiaalisiksi teemoiksi.
4. Teemojen tarkistaminen	Tarkistetaan toimivatko teemat koko aineistossa ja temaattisen kartan luominen.
5. Teemojen määrittely ja nimeäminen	Jatkuva teemojen hiominen, selkeiden nimien ja määritelmien luominen
6. Raportointi	Reflektointi tutkimuskysymyksiin ja kirjallisuuden, raportin kirjoittaminen

Vuonna 2017 Braun ja Clarke kommentoivat metodologiaa lisää. Heidän mukaansa temaattinen analyysi ei ole sidottu mihinkään teoreettiseen viitekehykseen, joten sitä voidaan käyttää joustavasti ja monipuolisesti. Joustavuus pätee myös tutkimuskysymyksiin, aineiston kokoon ja aineiston keruumenetelmiin.

Analyysia onkin käytetty esimerkiksi tunnistamaan esimerkillisiä verkko-opettajia (Edwards, Perry, & Janzen, 2011), moderaattoreiden aktiviteettien analysoimiseksi asynkronisessa verkkotukifoorumilla (Smedley & Coulson, 2017) ja sosiaalisen verkostanalyysin ohella vuorovaikutusverkkojen analysoimiseen verkkoroolipelin kiltayhteisössä (Ang, 2010).

Temaattinen analyysi on menetelmänä lähellä toista usein käytettyä laadullista menetelmää eli sisällönanalyysiä, jolla on myös vakiintuneempi tutkimusperinne (Vaismoradi, Turunen, & Bondas, 2013). Temaattista analyysiä ei ole välttämättä edes nimetty, jos sitä on käytetty. Lisäksi Braun ja Clarke (2006) väittävät, että tutkimuksissa temaattinen analyysi on saatettu nimetä diskurssi- tai sisällönanalyysiksi, vaikka todellisuudessa metodi olisikin lähempänä temaattista analyysiä.

Vaismoradi et. al (2013) tarkastelevat review artikkelissaan temaattista analyysiä ja sisällönanalyysiä. He esittävät, että temaattinen analyysi on selkeästi oma metodi, vaikka sillä onkin paljon samaa laadullisen sisällönanalyysin kanssa.

Temaattisessa analyysissä teemat voivat olla abstraktiivisempia ja subjektiivisempia kuin koodit sisällön analyysissä. Temaattisessa analyysissä ei juurikaan käytetä vertauskoodausta juuri teemojen subjektiivisuuden takia. Lisäksi teemojen olisi löydettävä koko aineistosta, mutta sisällön analyysissä vastaavat teemat saattavat esiintyä koodien frekvenssin myötä. Koodien frekvenssien laskeminen sisällön analyysissä mahdollistaa aineiston kvantitatiivisen tarkastelun, mikä ei sellaisenaan välttämättä ole mahdollista te-

maattisessa analyysissä.

Vaikka temaattisen analyysin pohjalta ei voida ainakaan suoraan tehdä kvantitatiivisia johtopäätöksiä keskustelusta, voidaan sen avulla varmistua keskusteluiden aiheista ja tarkastella, onko keskustelu ollut tehtävänannon mukaista.

## 3.2 Sosiaalinen verkostanalyysi

Sosiaalinen verkostanalyysi (Social Network Analysis, SNA) on määritelty joukoksi verkkoteoreettisia menetelmiä sosiaalisten verkostojen mallintamiseksi ja analysoimiseksi (Wasserman, 1994). Menetelmässä käytetään myös tilastollisia ja todennäköisyyksiin pohjautuvia malleja sekä sosiologiasta kehittyneitä teorioita. Tässä osiossa tutustutaan verkkoteorian perusteisiin sekä temporaalisiin verkkoihin. Lisäksi määritellään asynkroninen temporaalinen verkko. Lopuksi tutustutaan verkoissa esiintyviin rooleihin ja niiden mahdolliseen käyttöön SNA:ssa.

### 3.2.1 Verkkoteorian perusteita

Matematiikassa verkkoja tutkivaa alaa kutsutaan graafi- tai verkkoteoriaksi. Termeillä verkko ja graafi tarkoitetaan periaatteessa samaa, mutta verkko on intuitiivisempi ja modernimpi termi kuvaamaan konseptia (Estrada, 2012).

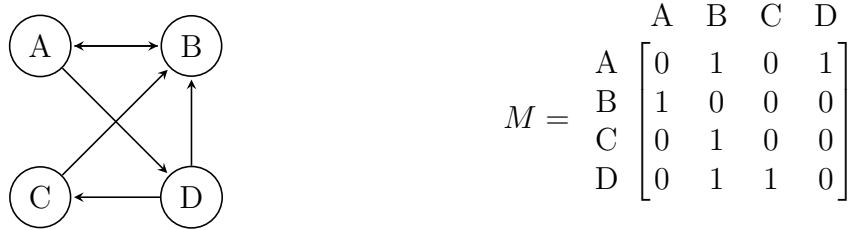
Yksinkertaisen määritelmän mukaan verkko tai graafi on kuvallinen esitys systeemistä, joka koostuu solmuista (Node) ja solmuja yhdistävistä linkeistä (Edge) (Estrada, 2012). Solmut ovat systeemin jäseniä ja linkit kuvaavat jäsenien välistä yhteyttä. Solmut voivat kuvata esimerkiksi ihmisiä, tietokoneita tai kaupunkeja, ja linkit vastaavasti sosiaalisia suhteita, verkkokaapeleita tai rautateitä. Jos verkko sisältää kaikki mahdolliset linkit solmujen välillä, sanotaan verkon olevan täydellinen.

Verkon kuvallisella esityksellä, eli solmujen sijainnilla ja linkkien pituudella ei ole yhteyttä fyysikaaliseen systeemiin, vaan se kuvaa vain ja ainoastaan systeemin jäsenien välisiä yhteyksiä. Tämä tarkoittaa käytännössä sitä, että jos kahdella verkolla on täysin samat solmut ja niiden välillä on samat linkit, ovat verkot samanlaisia eli ne ovat keskenään isomorfisia (Savolainen, 2001).

Verkko voi olla suuntaamaton tai suunnattu. Suunnatussa verkossa linkeillä on suunta, joka merkitään nuolilla. Suunnatussa verkossa voi siis kahden solmun välillä olla yksi tai kaksi linkkiä. Ollakseen isomorfeja, kahden suunnatun verkon linkkien on myös oltava samat koko verkossa.

Verkot määritellään formaalisti järjestettyjen parien joukkona s.e. jos kahden verkon alkion välillä on linkki, muodostavat ne parin. Verkkojen matemaattiseen käsittelyyn käytetään kuitenkin usein erilaisia matriisiesityksiä (Savolainen, 2001). Tässä työssä käy-

tetään matriisia, jossa kuvataan verkon solmujen väliset linkit. Käytetään tästä matriisista termiä viereisyysmatriisi (Adjacency matrix). Viereisyysmatriisissa  $A$  alkio  $a_{ij} = 1$ , jos solmujen  $i$  ja  $j$  välillä on linkki, muuten  $a_{ij} = 0$ . Viereisyysmatriisi on siis aina neliömatriisi, jonka päädiagonaalilla kaikki alkiot ovat nollia, kun ei solmulla ei voi olla linkkiä itsensä kanssa eli ns. luuppeja. Suuntaamattoman verkon viereisyysmatriisi on aina symmetrinen, mutta suunnatun verkon viereisyysmatriisissa  $a_{ij} \neq a_{ji}$ . Suunnatulle verkolle  $a_{ij} = 1$ , jos verkossa on linkki solmusta  $i$  solmuun  $j$ . Kuvassa 3.1 on esitetty yksinkertainen neljän solmun suunnattu verkko ja sitä vastaava viereisyysmatriisi.



**Kuva 3.1:** Solmujen A, B, C ja D muodostama yksinkertainen suunnattu verkko ja sitä vastaava viereisyysmatriisi  $M$ .

Viereisyysmatriisin avulla voidaan helposti laskea verkon keskeisyysmittoja, kuten esimerkiksi vastavuoroisten linkkien osuus. Keskeisyysmittojen perusteella voidaan analysoida verkkoja tai niiden ominaisuuksia (Zweig, 2016).

Verkkoa ja sen rakennetta voidaan tutkia myös tarkastelemalla sen osia eli ns. aliverkkoja. Tämä on erityisen hyödyllistä suurten tai monimutkaisten verkkojen tapauksessa. Verkko  $G'$  on verkon  $G$  aliverkko, jos  $G' \subseteq G$ . Aliverkko voi olla indusoitu, jolloin vaaditaan että aliverkko sisältää kaikki ne alkuperäiset linkit, jotka löytyvät verkosta. Aliverkko on osittainen tai vajaa, jos se ei sisällä kaikkia alkuperäisiä linkkejä.

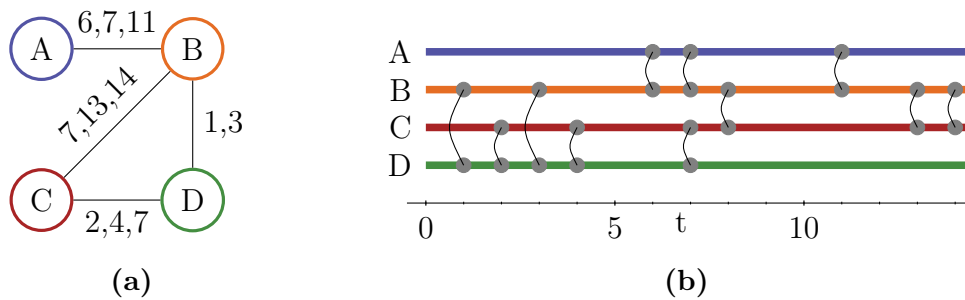
### 3.2.2 Temporaaliset verkot

Verkot kuvaavat usein dynaamista systeemiä ja Zweig (2016) esittääkin kirjassaan, että systeemin kuvaaminen verkkona on mielekästä vain jos solmujen välillä on jokin transitiivinen vuoro. Toisin sanoen linkkien pitää kuvastaa jonkin asian siirtymistä verkossa. Voidaan esimerkiksi muodostaa verkko, jossa solmuina on ihmisiä ja heidän välillään on linkki, jos molempien nimissä esiintyy sama kirjain. Kyseinen verkko on täysin määritelmän mukainen, mutta todellisuudessa henkilöiden välillä ei ole yhteyttä.

Vaikkakin verkoilla mallinnetaan dynaamista systeemiä, on niiden esitys käytännössä staattinen. Holme ja Saramäki (2012) ovat esittäneet temporaalisten verkkojen nousevan alan. Temporaalisissa verkoissa verkon linkkeihin lisätään aikamääre, joka kertoo milloin linkki on aktiivinen. Tämänlaisia verkkoja esiintyy esimerkiksi ihmisten välisessä kommunikaatiossa (Iribarren & Moro, 2009; Eckmann, Moses, & Sergi, 2004) ja fyysisis-

sä kontakteissa (Isella et al., 2011; Eagle & Pentland, 2006), solubiologiassa (Przytycka, Singh, & Slonim, 2010), neuroverkoissa (Valencia, Martinerie, Dupont, & Chavez, 2008) ja ekologisissa verkoissa (de Ruiter, Wolters, & Moore, 2005).

Tällä tavalla esitettynä linkit eivät muodosta järjestettyjen parien joukkoa, vaan ne muodostavat kolmikkoja, jossa jokaiseen pariin lisätään aikamääre. Esimerkiksi solmut A ja B muodostavat ajanhetkellä  $t$  linkin, joka esitetään kolmikkona  $(A,B,t)$ . Tällöin sama pari saattaa esiintyä joukossa useamman kerran eri aikamääreillä. Temporaalinen verkko voidaan visualisoida tavallisen verkon tapaan, mutta jokaiseen linkkiin lisätään näkyviin aikamääre, jolloin se on ollut aktiivinen. Toinen tapa esittää verkon temporaalinen rakenne on asettaa solmut ja linkit aikajanalalle. Molemmat visualisoinnit on esitetty kuvassa 3.2.



**Kuva 3.2:** Kaksi eri esitystapaa temporaaliselle verkolle. Ensimmäisessä ajankohdat, jolloin linkki on aktiivinen, on merkitty linkin oheen. Toisessa solmut ja linkit on asetettu aikajanalalle, josta nähdään, milloin verkon jäsenet ovat olleet linkitettyinä.

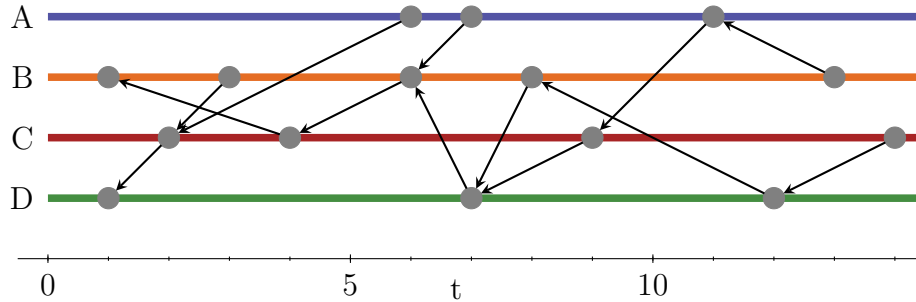
Temporaalisen verkon hyödyt ovat ilmeiset, kun tarkastellaan vaikkapa tiedon siirtymistä ihmisten muodostamassa ryhmässä. Otetaan esimerkiksi kuvan 3.2 verkko, joka nyt kuvaa neljän ihmisen välistä vuorovaikutusta. Tässä esimerkissä oleellinen tieto voi siirtyä vain, kun ihmiset ovat kontaktissa keskenään, joka vastaa sitä että verkon linkki on ollut aktiivinen. Oletetaan, että henkilöllä A on jokin merkittävä tieto ajan hetkellä  $t=0$ . Tämä tieto ei voi siirtyä henkilölle D vaikka A ja D ovatkin B:n välityksellä yhteydessä. Toisin päin tiedon siirtyminen kuitenkin on täysin mahdollista. Saman tyyppisillä temporaalisilla yhteyksillä on merkitystä esimerkiksi infektioiden leviämisen mallintamisessa (Vazquez, Rácz, Lukács, & Barabási, 2007).

### 3.2.3 Asynkroninen temporaalinen verkko

Edellä esitetty temporaalinen verkko ei sellaisenaan sovellu asynkronisen viestinnän esittämiseen. Asynkronisuuden takia linkkiin ei voida liittää vain yhtä aikamäärettä, jolla linkki olisi aktiivinen eikä sitä siten voida esittää järjestettynä kolmikkona. Esitys voidaan kuitenkin laajentaa asynkroniseksi lisäämällä linkkiin toinenkin aikamäärä. Kutsutaan tätä esitystä asynkroniseksi temporaaliseksi verkoksi, ATV.

Nyt jokainen linkki kuvataan nelikkona, jossa on solmujen lisäksi lähetys- ja vas-

taanottoaika. Asynkronisen verkon on oltava suunnattu, sillä yhteys on ajallisesti suuntautunut. Suunta voidaan kuitenkin valita tilanteen mukaan ja se voi kuvata viestien lähettämistä tai esimerkiksi tiedon kulkeutumista. Esimerkiksi A on lähettänyt viestin ajan hetkellä  $t_1$ , johon B on vastannut ajan hetkellä  $t_2$ . Tämä linkki voidaan kuvata nelikkona  $(A, B, t_1, t_2)$ .



**Kuva 3.3:** Asynkroninen temporaalinen verkko, jossa jokainen linkki solmujen välillä on tapahtuma jonkin ajanjakson yli.

Kuvassa 3.3 on visualisoituna yksinkertainen asynkroninen temporaalinen verkko, joka voisi esimerkiksi kuvata viestiketjuja foorumikeskustelussa. Jokainen solmu kuvaa viestiä, jonka ko. henkilö on lähettänyt. Aikajanalla nähdään suoraan myös viestin lähetysaika. Linkit solmujen välillä kuvaavat viestien välistä yhteyttä, eli sitä mihin viestiin on vastattu ja milloin.

### 3.3 Roolit

McDonnell et. al (2014) ovat esittäneet miten verkosta voidaan laskea erilaisia rooleja verkon jäsenille viereisyysmatriisin avulla. Roolit löytyvät kolmen solmun muodostamisesta, triadisista motiffeista eli aliverkon kuvista. Nämä motifit ovat määritelty sellaisiksi aliverkoiksi, jotka esiintyvät useammin kuin satunnaisessa verkossa voisi olettaa.

Satunnaisen verkon malli (Random Graph Model, RNG) on tapa tuottaa satunnaisia verkkoja eri parametreilla. Vertaamalla tarkastelun kohteena olevaa verkkoa satunnaiseen verkkoon, voidaan tehdä johtopäätelmiä onko verkon rakenne muutakin kuin vain satunnainen. Kun verkko on pieni, ei kuitenkaan voida suoraan käyttää RNG:tä analysointityökaluna (Zweig, 2016). Roolit voidaan kuitenkin laskea pienellekin verkolle.

Verkossa voi olla erilaisia kolmen solmun rooleja 30, jotka jaetaan yhdeksään eri roolityyppiin. Kaikki roolit esiintyvät 13 erilaisessa aliverkon kuvissa. Osassa aliverkon kuvissa siis esiintyy sama rooli useammin. Kuvassa 3.4 on esitetty kaikki roolit, niiden tyypit sekä laskukaavat viereisyysmatriisista.


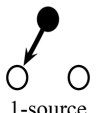

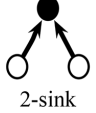
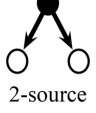



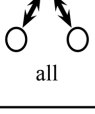
McDonnell et. al jakavat roolit funktionaalisiin ja rakenteellisiin, joita vastaavat verkon osittaiset ja indusoidut aliverkot. Rakenteelliset roolit kuvaavat ikään kuin verkon



anatomisia rakenneosia, kun taas funktionaaliset roolit kuvastavat mahdollisia vuorovaikutusmalleja verkossa. Vertaa esimerkiksi neuroverkon synapseja ja mahdollisia synaptisia aktivointikuvioita.

Yhdellä solmulla voi olla siis vain yksi rakenteellinen rooli kolmikossa, mutta funktionaalisia voi olla useampikin, sillä esimerkiksi rakenteellisessa roolityypillä 3 esiintyy myös funktionaaliset roolityypit 1 ja 2. Tosin sanoen rakenteellinen vastavuoroisuus voi esiintyä funktionaalisti myös lähteenä tai nieluna. Vastaavasti rakenteellisella roolityypillä 9 esiintyy kaikki muut roolityypit funktionaalisina.

Ensimmäinen roolityyppi on vastaanottaja tai nielu ja toinen vastaavasti lähde. Kolmas roolityyppi on vastavuoroinen, joissa linkki kahden solmun välillä on molemminpuolinen, mikä suhteena kuvastaa vastavuoroisuutta. Kolmas ja neljäs roolityyppi ovat samantyyllisiä kuin ensimmäinen ja toinen, mutta roolin solmulla on samanlainen yhteys kahteen muuhun solmuun. Kuudes roolityyppi eroaa aiemmista, sillä se on ikään kuin välittäjätyyppinen, jossa roolin solmu toimii esimerkiksi tiedon tai vuorovaikutuksen välittäjänä kahden muun solmun välillä, joilla ei ole omaa yhteyttä. Seitsemäs ja kahdeksas ovat yhdistelmiä näistä aiemmista ja yhdeksäs roolityyppi pitää sisällään täysin vastavuoroiset yhteydet.

Role	Motif-roles			
 1-sink	$r = 1$ $m = 1$ $d_r = 2$ $\mathcal{A}^\top (\mathcal{A} - \mathcal{I}) \mathbf{1}$	$r = 2$ $m = 2$ $d_r = 1$ $((\mathcal{A}^\top)^2 - \mathcal{R}) \mathbf{1}$	$r = 3$ $m = 3$ $d_r = 1$ $(\mathcal{A}^\top \mathcal{R} - \mathcal{R} \circ \mathcal{A}) \mathbf{1}$	
 1-source	$r = 4$ $m = 2$ $d_r = 1$ $(\mathcal{A}^2 - \mathcal{R}) \mathbf{1}$	$r = 5$ $m = 4$ $d_r = 2$ $\mathcal{A} (\mathcal{A}^\top - \mathcal{I}) \mathbf{1}$	$r = 6$ $m = 7$ $d_r = 1$ $(\mathcal{A} \mathcal{R} - \mathcal{A} \circ \mathcal{R}) \mathbf{1}$	
 1-recip	$r = 7$ $m = 3$ $d_r = 1$ $(\mathcal{R} \mathcal{A} - \mathcal{R} \circ \mathcal{A}) \mathbf{1}$	$r = 8$ $m = 7$ $d_r = 1$ $(\mathcal{R} \mathcal{A}^\top - \mathcal{R} \circ \mathcal{A}) \mathbf{1}$	$r = 9$ $m = 8$ $d_r = 2$ $\mathcal{R} (\mathcal{R} - \mathcal{I}) \mathbf{1}$	
 2-sink	$r = 10$ $m = 4$ $d_r = 1$ $\frac{1}{2} \mathcal{A}^\top \mathbf{1} \circ (\mathcal{A}^\top - \mathcal{I}) \mathbf{1}$	$r = 11$ $m = 5$ $d_r = 1$ $((\mathcal{A}^\top \mathcal{A}) \circ \mathcal{A}^\top) \mathbf{1}$	$r = 12$ $m = 6$ $d_r = 1$ $\frac{1}{2} ((\mathcal{A}^\top \mathcal{R}) \circ \mathcal{A}^\top) \mathbf{1}$	
 2-source	$r = 13$ $m = 1$ $d_r = 1$ $\frac{1}{2} \mathcal{A} \mathbf{1} \circ (\mathcal{A} - \mathcal{I}) \mathbf{1}$	$r = 14$ $m = 5$ $d_r = 1$ $(\mathcal{A}^2 \circ \mathcal{A}) \mathbf{1}$	$r = 15$ $m = 11$ $d_r = 1$ $\frac{1}{2} ((\mathcal{A} \mathcal{R}) \circ \mathcal{A}) \mathbf{1}$	
 relay	$r = 16$ $m = 2$ $d_r = 1$ $\mathcal{A}^\top \mathbf{1} \circ \mathcal{A} \mathbf{1} - \mathcal{R} \mathbf{1}$	$r = 17$ $m = 5$ $d_r = 1$ $((\mathcal{A} \mathcal{A}^\top) \circ \mathcal{A}^\top) \mathbf{1}$	$r = 18$ $m = 9$ $d_r = 3$ $(\mathcal{A}^2 \circ \mathcal{A}^\top) \mathbf{1}$	$r = 19$ $m = 10$ $d_r = 1$ $((\mathcal{A} \mathcal{R}) \circ \mathcal{A}^\top) \mathbf{1}$
 relay & sink	$r = 20$ $m = 7$ $d_r = 1$ $(\mathcal{R} \mathbf{1}) \circ ((\mathcal{A}^\top - \mathcal{I}) \mathbf{1})$	$r = 21$ $m = 10$ $d_r = 1$ $((\mathcal{R} \mathcal{A}) \circ \mathcal{A}^\top) \mathbf{1}$	$r = 22$ $m = 11$ $d_r = 2$ $((\mathcal{A}^\top \mathcal{A}) \circ \mathcal{R}) \mathbf{1}$	$r = 23$ $m = 12$ $d_r = 1$ $(\mathcal{R}^2 \circ \mathcal{A}^\top) \mathbf{1}$
 relay & source	$r = 24$ $m = 3$ $d_r = 1$ $(\mathcal{R} \mathbf{1}) \circ ((\mathcal{A} - \mathcal{I}) \mathbf{1})$	$r = 25$ $m = 6$ $d_r = 2$ $((\mathcal{R} \mathcal{A}) \circ \mathcal{A}) \mathbf{1}$	$r = 26$ $m = 10$ $d_r = 1$ $(\mathcal{A}^2 \circ \mathcal{R}) \mathbf{1}$	$r = 27$ $m = 12$ $d_r = 1$ $((\mathcal{A} \mathcal{R}) \circ \mathcal{R}) \mathbf{1}$
 all	$r = 28$ $m = 8$ $d_r = 1$ $\frac{1}{2} \mathcal{R} \mathbf{1} \circ ((\mathcal{R} - \mathcal{I}) \mathbf{1})$	$r = 29$ $m = 12$ $d_r = 1$ $((\mathcal{R} \mathcal{A}) \circ \mathcal{R}) \mathbf{1}$	$r = 30$ $m = 13$ $d_r = 3$ $\frac{1}{2} (\mathcal{R}^2 \circ \mathcal{R}) \mathbf{1}$	

**Kuva 3.4:** McDonnelin et. al (2014) tuottama kuva, jossa on esitetty roolit  $r$ , missä verkon alikuvassa ne esiintyvät  $m$ , kuinka monta ko. roolia verkon alikuvassa on  $d_r$  ja alimmaisena funktionaalisia rooleja vastaavat laskukaavat, joissa  $\circ$  osoittaa hadamadrin tuloa,  $\mathcal{A}$  on viereisyysmatriisi,  $\mathbf{1}$  on  $\mathcal{N} \times 1$  yksikkö-pystyvektori,  $\mathcal{I}$  on  $\mathcal{N} \times \mathcal{N}$  yksikkömatriisi ja  $\mathcal{R} = \mathcal{A} \circ \mathcal{A}^\top$ .

Rooleja tai rakenteita neljän tai useamman solmun aliverkolle ei ole määritetty. Tämä johtuu siitä, että jo neljällä solmulla on 199 suunnattua aliverkkoa ja rooleja voisi mahdollisesti olla jopa  $4 \cdot 199 = 796$ . Toisaalta kolmen solmun aliverkon rooleilla voidaan analysoida verkon ominaisuuksia. Roolien avulla voidaan tunnistaa esimerkiksi solmu, jol-

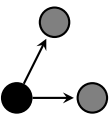
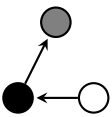
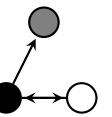
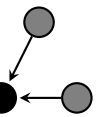
la on keskeinen asema verkossa. Rooleja voidaan suoraan verrata joihinkin keskeisyysmittoihin.

Tässä työssä halutaan tarkastella minkä tyyppisiä rooleja keskusteluissa esiintyy. Verkkokeskustelussa rakenteelliset roolit kuvaavat keskustelun teknistä rakennetta, mutta funktionaaliset roolit taas keskustelun taustalla vaikuttavia mahdollisia sosiaalisia rakenteita. Jos viestiketjuista tarkastellaan sosiaalisia rakenteita ja käytetään verkon suuntana viestin lähetyssuuntaa, kuvastaa rooli yksi esimerkiksi arvostuksesta nauttimista ja vastaavasti lähde on arvostuksen osoittaja. Relay-tyypin roolit kuvastavat taas kaverin kaveri-tyyppisiä suhteita, joissa arvostus välittyy erillisen henkilön kautta. Molempiin suuntaan oleva linkki taas kuvastaisi mahdollista vastavuoroista vertaisarvostusta.

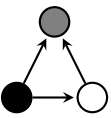
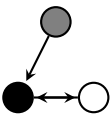
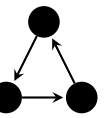
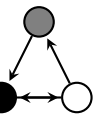
Roolit voidaan laskea käyttäen McDonnellin esittämiä kaavoja, jotka ovat kuvassa 3.4. Sijoittamalla keskustelusta muodostettu viereisyysmatriisi roolin kaavaan, saadaan tuloksena vektori, jossa on lukumäärä monessako ko. roolissa solmu on. Laskemalla yhteen kyseisen roolityyppiin liittyvien roolien lukumäärät, saadaan suoraan selville, montako kyseisen tyyppin roolia verkossa yhteensä voi olla. Tässä halutaan tarkastella vain yhdeksää roolityyppiä kaikkien kolmenkymmenen erillisen roolin sijaan. Tämä johtuu yksinkertaisesti siitä, että pienessä verkossa ei riitä resoluutio kaikkiin rooleihin.

Taulukossa 3.2 on esitetty ne aliverkkojen kuvat ja roolit, jotka esiintyvät kuvan 3.1 esimerkkgraafissa.

**Taulukko 3.2:** Esimerkkgraafin roolit, aliverkon kuvat, joissa roolit esiintyvät, sekä kuinka monta ko. aliverkon kuvaa graafista muodostuu.

Aliverkon kuva	1	2	3	4
				
lkm	2	6	1	3
roolit	1, 13	2, 4, 16	3, 7, 24	5, 10

---

Aliverkon kuva	5	7	9	10
				
lkm	2	2	1	1
roolit	11, 14, 17	6, 8, 20	18	19, 21, 26

Taulukossa 3.3 on laskettu esimerkkgraafin ja täydellisen graafin roolisummat. Verkossa esiintyvien roolien lukumäärä voidaan normalisoida täydellisen graafin roolisummien avulla. Koska täydellisen graafin roolisummat ovat suurimmat mahdolliset, saadaan

esimerkkiverkon roolien osuus maksimista jakamalla sen roolisummat täydellisen verkon vastaavilla. Taulukon 3.3 alimmalla rivillä on normitetut roolisummat esimerkkigraafille.

**Taulukko 3.3:** Esimerkkigraafin sekä täydellisen neljän solmun graafin roolien summat.

rooli	1	2	3	4	5	6	7	8	9
esim. verkko	11	14	3	5	4	12	3	2	0
täyd. verkko	72	72	72	48	48	96	96	96	48
normitettuna	0,15	0,19	0,04	0,10	0,08	0,13	0,03	0,02	0,00

Tällä tavoin laskemalla verkossa esiintyvät funktionaaliset roolit tai niiden osuus maksimimäärästä, voidaan saada tietoa esimerkiksi verkossa esiintyvien vuorovaikutusten luonteesta. Esimerkiksi edellä esitettyjen esimerkkigraafin normitettujen tulosten avulla on helppo nähdä, että ko. verkko sisältää vain vähän resiprokaalisia eli vastavuoroisia rooleja (3, 7-9) ja vastaavasti suurin osuus on rooleilla, joissa vuorovaikutus on yksisuuntaista.

Kyseissä yksinkertaisessa verkossa tämä tulos on helposti nähtävillä suoraan, sillä ainoastaan solmujen A ja B välinen linkki on kaksisuuntainen. Verkossa, jossa on enemmän vuorovaikutuksia rooleja ei ole mahdollista tarkastella yhtä suoraviivaisesti. Vastaavasti temporaalisissa verkoissa rooleja tai niiden ajallista kehittymistä on mahdoton tarkastella yksinkertaisesti visuaalisesti.

### 3.3.1 Roolien yhteys sosiaaliin verkkoihin

McDonnell et. al käyttävät rooleja jonkin asian välittäjinä. Esimerkissään he ovat tarkastelleet signaalien välittymistä alkueläimen neuroverkossa. Sosiaalisia suhteita kuvaavissa verkoissa yhteys kuvataan usein esimerkiksi arvostuksen tai luottamisen kautta. Tällöin suunta on arvostettua henkilöä kohti.

Jos oletetaan, että arvostetun henkilön viestiin vastataan herkemmin, tällöin verkossa viestien suunta vastaa sosiaalisen verkon suuntaa. Jos taas halutaan tarkastella tiedon siirtymistä tai yhteistä tiedonrakentumista keskustelussa, tulisi verkon suunta olla käänteinen viestien suunnan kanssa. Tällainen käänteinen rakenne vastaisi myös paremmin artikkelissa esitettyjä nimityksiä roolityypeille: nielu, lähde yms. Onneksi laskennallisesti tilanne on yksinkertainen, sillä käänteinen verkko saadaan transponoimalla viereisyysmatriisi.

Tarkastellaan tässä nimenomaan sosiaaliin rakenteisiin kytkeytyviä rooleja. Edellä on näytetty miten verkosta saadaan laskettua funktionaaliset roolit, joita on selkeästi paljon enemmän kuin rakenteellisia rooleja. Herää kenties kysymys miksi tarkastellaan funktionaalisia rooleja todellisten rakenteiden sijaan. On kuitenkin syytä muistaa, että

verkkokeskustelun rakenne ei millään perusteella vastaa suoraan ryhmän sosiaalisia rakenteita. Funktionaalisilla rooleilla saadaan esille kaikki ne mahdolliset rakenteet, jotka voivat esiintyä sosiaalisessa verkossa. Funktionaalisia rooleja tai niiden jakautumista voidaan siten verrata teorioihin sosiaalisista rakenteista ja niiden muodostumisesta.

Lawlerin (et. al 2000, 2001, et. al 2008) *Affect theory of social Exchange* kuvaa kuinka tunteet, positiiviset ja negatiiviset, säätelevät yhteyksien muodostumista ihmisten välillä. Teorian mukaan toistuvat vuorovaikutukset synnyttävät ns. sosiaalisia mikrojärjestyksiä (Micro Social Order). Lawler esittää neljä erilaista vuorovaikutustyyppiä, joista hyödyllisin on produktiivinen. Siinä ryhmän jäsenet yhdessä tuottavat kollektiivisen hyödykkeen ja jokainen jäsen toimittaa sekä saa hyötyä ryhmältä. Mallissa ei suoraan ole edellä esitettyjen tyyppisiä rooleja, mutta se kuvaa vastavuoroisuuden edulliseksi ryhmän jäsenille. Positiivisten kokemusten kautta ryhmän koheesio kasvaa ja ryhmän väliset suhteet kehittyvät resiprokaalisiksi.

Skvoretz (1996) on kuvannut miten ryhmän sisäinen arvostusasema- eli statusjärjestys kehittyy pienissä ryhmissä. Tässä mallissa sivustajakatsojilla on merkittävä rooli, sillä vuorovaikutuksen havaitseminen synnyttää osaltaan rakenteita verkkoon, joka kuvaa statusjärjestystä. Skvoretz on kuvannut vakaita triadeja, joita vastaavat McDonnellin motiffeja 5, 9, 11, jossa vastavuoroinen yhteys konfliktina sekä 6 ja 13, joissa vastavuoroinen yhteys kuvaa molemmin puoleista kunnioitusta. Tässä vakaa tarkoittaa, että muiden agenttien käyttäytyminen ei voi häiritä kyseisiä rakenteita. Nämä rakenteet kuvastavat johtaja-tyyppisiä rakenteita, joka verkkokeskusteluissa näyttäytyisi tyyppin neljä rooleina, joissa kaksi jäsentä kunnioittavat ja siten viestivät kolmannen kanssa.

Mallin mukaan järjestys stabiloituu nopeasti, jos ryhmällä on jo etukäteen jonkinlainen sisäinen järjestys. Lisäksi alussa aktiivisesti vuorovaikutuksiin osallistuneet, ovat aktiivisia myös myöhemmin sillä heillä on korkeampi asema ryhmän sisällä. Huomionarvoista on myös, että mallissa tarvitaan vähintään 22 vuorovaikutusta, että statusjärjestys on valmis.

Lusher (2014) on puolestaan tutkinut luottamussuhteiden rakennetta pienryhmissä. Hänen mallissaan on 13 erityyppistä vuorovaikutusrakennetta, jotka jaetaan luottamustageneroiviin ja -rajoittaviin prosesseihin. Luottamusta lisääviä rakenteita ovat resiprokaaliset dyadit ja niin kutsutut ystävän ystävä triadit, joka vastaa roolia 17. Luottamusta rajoittaa jaettu suosio, joka rakenteena vastaa roolia 10. Lusherin mukaan ryhmissä, joissa on paljon luottamustageneroivia rakenteita, menestyvät paremmin.

Edellä esitettyjen mallien rakenteet eivät suoraan vastaa McDonnellin rooleja, mutta niiden perusteella voidaan kuitenkin väittää, että verkkokeskusteluissa pitäisi esiintyä tyyppien 3, 7, 8 ja 9 mukaisia rooleja, jos vuorovaikutuksia on vain ollut tarpeeksi. Lisäksi syklisiä rooleja eli tyyppin 6 ei pitäisi juurikaan esiintyä (Yoon, Thye, & Lawler, 2013), ja tasavertaisissa ryhmässä pitäisi nopeasti esiintyä resiprokaalisia rooleja.

Tätä vahvistaa myös Koposen & Nousiaisen (2016) kehittämä laskennallinen malli vastavuoroisten arvostusrakenteiden syntymiselle. Malli esittää minkälaisia dyadisia ja triadisia rakenteita syntyy riippuen ryhmän jäsenten välisestä arvostuksesta sekä jokaisen itsearvostuksesta. Malli ennustaa transitiota tasa-arvoisista triadeista vahvoihin dyadeihin, kun jäsenten herkkyys kilpailulliseen vertailuun lisääntyy.

Jos verkkokeskusteluissa ilmenee saman tyyppisiä rakenteita kuin sosiaalisen vuorovaikutuksen malleissa, antaa se osviittaa näiden väliseen mahdolliseen yhteyteen. Toisaalta se luo pohjaa roolien laskemisen käytölle sosiaalisten vuorovaikutusten tutkimiseen, mutta ennen kaikkea se vahvistaa menetelmää sosiaalisten suhteiden tarkasteluun.

## 4. Tutkimus

### 4.1 Tutkimuskysymykset

Tämän työn pääasiallinen tavoite on löytää menetelmät ja työkalut verkkokeskustelun auki purkamiseen ja analysoimiseen. Käytännössä tämä jakaantuu muutamaaan eri osaan ja vaiheeseen. Ensinnäkin pitää olla selvillä mistä verkkokeskustelussa keskusteltiin, jotta tätä seuraavilla analyyseillä olisi järkevä perusta. Toisaalta myös keskustelun sisältö on hyödyllistä tietää, jos halutaan arvioida keskustelua kurssisuorituksena.

Usein myös keskustelun rakenteen hahmottaminen nopeasti on arvokasta. Keskustelun rakennetta voi olla haastavaa seurata suoraan tarkastelemalla keskustelua foorumi-alustalla ja etenkin asynkronisuuden hahmotus suoraan on erittäin hankalaa. Tavoitteena onkin keksiä esitystapa keskustelun rakenteelle, joka on helposti hahmotettavissa nopealla silmäilyllä.

Keskustelun rakenteesta olisi tarkoitus selvittää myös keskustelussa esiintyneet roolit. Lisäksi roolein mahdollinen aikakehitys olisi hyödyllistä saada esille. Tähänkin halutaan löytää sopiva graafinen esitys.

Tavoitteiden pohjalta muodostetut tutkimuskysymykset ovat:

1. Mistä keskusteltiin?
2. Voidaanko keskustelu esittää selkeästi graafisesti?
3. Minkälaisia rooleja keskustelussa esiintyy, mikä on niiden aikakehitys ja voidaanko ne esittää graafisesti?

### 4.2 Aineisto

Aineisto kerättiin aineenopettajien *Fysiikan käsitteenmuodostus II* -kurssilla käydyistä verkkokeskusteluista. Kurssi on suunnattu fysiikkaa toisena opetettavana aineena opiskeleville ja pääsääntöisesti kurssin osallistujat opiskelivat pääaineenaan matematiikkaa.

Opiskelijat tutustuivat kurssilla kvanttimekaniikan syntymiseen ja kehittymiseen perehtymällä aiheeseen liittyviin tieteellisiin artikkeleihin. Artikkeleiden pohjalta opiskelijoi-

den tuli tehdä argumenttikaaviot ja edelleen niiden avulla didaktiset rekonstruktiot.

Argumenttikaavioiden avulla selvitetään miten kvanttimekaniikan käsitteet on aikaan perusteltu tiedeyhteisössä. Argumenttikaavioiden aiheet ja niiden pohjalla käytetyt artikkelit on esitetty taulukossa 4.1. Argumenttikaavioissa on tarkoitus graafisesti kuvata tieteellisen argumentin rakenne. Argumentointi aloitetaan motiivista tai tavoitteesta ja etenee teorian, mahdollisen kokeen ja datamallin kautta johtopäätöksiin ja seuraamuksiin.

**Taulukko 4.1:** Argumenttikaavioiden aiheet ja niiden pohjalla käytettävät artikkelit

#	Aihe	Artikkeli
1.	Valosähköinen ilmiö ja energian kvantittuminen	Millikan R.A. <i>A direct photoelectric determination of Planck's <math>h</math>.</i>
2.	Comptonin ilmiö ja liikemäärän kvantittuminen	Compton A. H. <i>A quantum theory of scattering of x-rays by light elements</i> ja <i>The spectrum of scattered x-rays</i>
3.	Kaksoisrakokoe fotonille ja aaltohiukkasdualismi: valon hiukkasluonne	Rueckner W. & Titcomb P. <i>A lecture demonstration of single photon interference</i>
4.	Kaksoisrakokoe elektronille ja aaltohiukkasdualismi: elektronien aaltoluonne	Hobson A. <i>There are no particles, there are only fields</i>

Argumenttikaavioiden työstämiseen käytettiin avuksi verkkokeskusteluja kurssikohdaisella Moodle-alustalla. Keskusteluissa oli tarkoitus vertailla tehtyjä argumenttikaavioita sekä pohtia, miten argumenttikaavio voidaan laajentaa didaktiseksi rekonstruktioksi.

Opiskelijoille annettiin seuraavat ohjeet keskusteluihin liittyen:

*”Ensimmäinen tehtävä on liittää oma kaavio keskustelualueelle, jotta ryhmän jäsenet näkevät toistensa työt. Kaikkien kaaviot pohjautuvat samaan pohjamateriaaliin, joten tarkoitus onkin pohtia: Mitä yhteistä argumenttikaavioissanne on ja miten ne eroavat toisistaan? Millaista lähestymistapaa argumenttikaaviot heijastelevat? Miten argumenttikaaviota voi hyödyntää didaktisen rekonstruktion laatimisessa? Voiko jotain poistaa? Mitä pitää lisätä? Miten ja missä järjestyksessä asia tulisi opettaa?”*

Kurssille osallistui 16 opiskelijaa, jotka jakaantuivat neljän hengen keskusteluryhmiin. Neljän aiheen ja neljän ryhmän kautta analysoitavaksi saatiin yhteensä 16 eri verkkokeskustelua. Verkkokeskustelut oli jouduttu kopioimaan taulukkolaskentatiedostoihin, sillä kurssin Moodle-alue oli poistettu käytöstä. Kopioitaessa oli säilytetty viestien lisäksi keskustelun puurakenne eli kuka oli vastannut kenellekin ja viestien lähetysajat. Lisäksi opiskelijoiden nimet oli korvattu nimimerkeillä anonyymiteetin säilyttämiseksi. Kuvassan 4.1 on esimerkki aineistosta.



	A	B	C	D	E	F	G	H	I	J
1	14:03	A1:	Tässä minun argumenttikaavioni.							
2	14:04	A2:	Tässä on minun fotonien kaksoisrakokokeeseen liittyvä argumenttikaavioni.							
3	14:08	A4:	Tässä tuotokseni.							
4			14:30 A4:	Pääpiirteittäin kaaviot ovat varsin samankaltaisia. Seuraamuksena A1 on poiminut valon kvanttiluonteen A2stä ja						
5				14:39 A2:	Itse mietin myös tuota kvantittumista. Vaikka energia siirtyy elektronille kvantittuneena, ei tämä art					
6				14:41 A2:	En kuitenkaan ole vielä kunnolla tutustunut lisälukemisen (Kiddin artikkeli). Ehkä sen lukeminen se					
7				A1:	Aivan. Otin sen seuraamukseksi tämän artikkelin pohjalta tehdessä, sillä artikkelin Comments koh					
8				14:50	Onhan A2lläkin se tuolla mainittuna, vasta-argumentin kiistäminen-kohdassa.					
9					14:54 A2:	Joo, kyllä mainitsin sen siellä (vasta-argumenteissa). Se oli kuitenkin selkeästi mair				
10					15:00 A1:	Niin totta, minunkin tarvitsee ottaa tästä lisää selvää.				
11					16:12 A4:	Ja siis minullakin on tämä samassa kohtaa, kiistämisessä. Ottaisikohan luennoijan				
12					14:53 A1:	En ole vielä ehtinyt lukea lisääaineistoa. Tarvitsee lukea, jotta tämä kohta selkenee itselle.				
13					14:59 A3:	En ole vielä ehtinyt lukea lisälukemistoa, mutta tämä kohta tosiaan on hieman epäselvä. Itse aluks				
14					15:01 A3:	Siis että tämä hiukkasmaisuus havaitaan myös tässä kaksoisrakokokeessa, ja siter				
15					15:02 A2:	Taidat olla oikeassa. Näin minäkin sen ymmärrän.				
16					15:11 A1:	Olen samaan mieltä myöskin.				
					A1:	Tuli vielä mieleen tästä artikkelissa mainitusta valon kvanttiluonteen osoituksesta: ehkä sillä tarkoi				
						Tämä tuli mieleeni, kun luin asken Galilei 8:n duaalisen mallin selityksen alun: "Duaalisessa mallissa t				
						Ehkä artikkelissa mainitulla "quantum nature of light" tarkoitetaan juuri tätä, että missä mielessä hiukk				
17	14:15	A1:	Motiivit, tavoitteet ja todisteet ovat meillä kolmella pitkälti sisällyttäen samat.							
18	14:23	A2:	Minulla ja A4lla oli hieman erilaiset motiivit. Minä käsittelin vasta tavoitteessa tuon ongelman pohjista, kuinka y							
19	14:33	A1:	Aivan. Jädellakin on tuo pohdinta fotonien interferoinnista itsensä/toisten fotonien kanssa tavoitte							

Kuva 4.1: Alkuperäinen aineisto taulukoituna

Toisin sanoen analyysiä varten saatu aineisto on ns. sekundääristä eikä primääri-lähdettä ollut käytössä. Samasta syystä keskustelun metadattaa ei ole saatavilla. Tämä aiheuttaa mahdollisesti ongelmia. Jos aineistossa on esimerkiksi epäselvyyksiä, asiaa ei voida tarkistaa alkuperäisestä keskustelusta.

Muutamassa viestissä lähetysaika ja se, kenelle viesti on ollut vastaus, ovat olleet ristiriidassa. Tämä johtuu mitä ilmeisemmin siitä, että viestin kirjoittaja on jälkikäteen muokannut viestiään ja tuo muokkaus aika näkyy viestin lähetysaikana. Lisäksi ainakin yhdessä viestissä sisältö on selkeästi väärä. Viesti on keskustelun lopussa, mutta sen sisältö täysin sama kuin saman henkilön aloitusviestissä ko. keskustelussa.

### 4.3 Toteutus

Tutkimuksen käytännön osuus toteutettiin kahdessa osassa. Ensimmäiseksi saatu aineisto jaenneltiin uudestaan temaattista analyysiä varten. Temaattisen analyysin yhteydessä varmistettiin keskustelun rakenne, jonka lisäksi keskustelu jaettiin selkeisiin keskusteluepisodeihin. Teemat, viestien välinen yhteys ja episodit merkittiin erilliseen taulukkoon, jota käytettiin toisessa osassa.

Toisessa osassa kirjoitettiin Python-skriptit aineiston laskennalliseen käsittelyyn. Skripteillä visualisoitiin dialogirakenne temporaalisena asynkronisena verkostona, laskettiin roolien lukumäärät ja esitettiin roolisummien aika kehitys matriisina. Lisäksi oma skripti kirjoitettiin teemojen lukumäärän laskemiseksi. Käytetyt skriptit on esitetty liitteessä A.



mista. Lopulliset teemat, jotka ovat nähtävillä kuvauksineen ja esimerkkeineen taulukossa 4.3, saatiin kun teemoja hienosäädettiin vielä kerran analyysin jälkeen.

Teemojen löytymisen jälkeen jokaisen keskustelun viestit analysoitiin ja viesteistä löytyneet teemat kirjattiin ylös. Samanaikaisesti tarkastettiin keskustelun rakenne teemojen ja puurakenteen avulla. Tämä jouduttiin tekemään, koska osassa keskusteluista viestien sisältö ja puurakenne olivat ristiriitaisia. Tämä johtui ilmeisesti siitä, että opiskelijat eivät olleet tottuneet käyttämään Moodle-alustaa verkkokeskusteluihin ja vastasivat väärille henkilöille. Nämä epäselvyydet olivat kuitenkin käytännössä aina selvitettävissä viestien sisältöjen avulla. Keskustelun rakenne merkattiin yksinkertaisesti nuolilla tulostettuihin taulukoihin.

Keskustelun rakenteen ja viestien sisällön eli teemojen avulla keskustelun viestit jaettiin keskusteluepisodeihin. Episodi on sellainen selkeä temaattisesti eheä kokonaisuus, jossa opiskelijat ovat vuorovaikuttaneet keskenään eli viestejä on useampi kuin kaksi, ja viestien sisältö muodostaa yhtenäisen jatkumon. Toisin sanoen opiskelijat puhuvat yhdessä episodissa pääsääntöisesti yhdestä asiasta. Viestien laajuus huomioon ottaen teemoissa saattaa olla pientä vaihtelua, esimerkiksi episodissa keskustellaan teeman 1 alateemoista ja välillä viesteissä ilmenee myös teema 4.

Ensimmäisen analyysin jälkeen myös episodijaosta keskusteltiin ohjaajien kanssa ja episodeja päätettiin yhdistellä mm. siten, että yhdellä episodilla voi olla useampi aloitus. Näin episodien määrä väheni selkeästi, mutta samalla se toi lisää selkeyttä keskustelun rakenteeseen. Jako episodeihin ei siis perustu selkeään teoreettiseen malliin, vaan on enemminkin tapa jäsentää keskustelun rakennetta asiayhteyteen sopivalla tavalla. Jos keskustelun ohjeistus olisi ollut strukturoidumpi, voisi episodi olla tietyn kysymyksen tai tehtävänannon alla tapahtuva keskustelu.

Jokainen viesti taulukoitiin taulukkolaskentaohjelmaan siten, että jokaisen viestin kohdalle merkittiin, mihin viestiin se oli vastaus, viestin teemat ja episodi. Viestit identifioitiin nimimerkin ja lähetysajan perusteella. Jos viesti oli aloitusviesti, vastaus itselle tai se ei selkeästi vastannut kenellekään, merkittiin vastauksen kohde, eli kenelle, olevan itse viesti. Taulukossa 4.4 on esitetty mahdolliset tavat merkitä viestit. Taulukon ensimmäinen sarake on juokseva numerointi, joka vastaa keskustelun vastauksia. Negatiivisella numerolla on merkattu viestit, jotka eivät ole vastauksia kenellekään.



**Taulukko 4.2:** Suppea esimerkki miten viestit taulukoitiin

#	Aika1	Kenelle	Aika2	Kuka	Episodi	Teema
-1	1	D1	1	D1	0	0
-2	1	D4	1	D4	1	0
1	1	D4	7	D2	1	1b
1	1	D4	7	D2	1	4b
-3	2	D2	2	D2	0	0
2	3	D1	5	D2	0	7
3	3	D3	4	D2	2	5
4	4	D2	6	D3	2	5
-7	7	D4	7	D4	1	1a
5	7	D4	21	D2	1	1a
6	7	D4	9	D2	1	1d
7	7	D4	8	D3	1	1a
7	7	D4	8	D3	1	1d
8	7	D2	11	D1	1	3b
9	7	D2	10	D3	1	1e
9	7	D2	10	D3	1	4b
9	7	D2	10	D3	1	4c
10	8	D3	11	D4	1	1b
10	8	D3	11	D4	1	1d
10	8	D3	11	D4	1	1e

**Taulukko 4.4:** Miten erilaiset viestit taulukoitiin.

Tyyppi	#	Aika1	Kenelle	Aika2	Kuka	Episodi	Teema
Aloitusviesti	-1	1	A	1	A	1	0 / 1
Yksittäinen viesti	-1	1	A	1	A	0	0 / 1
Vastaus	1	1	A	2	B	1	1
Vastaus ilman episodua	1	1	A	2	B	0	1

Jos viesti sisälsi useita teemoja, tehtiin viestistä niin monta riviä kuin se sisälsi teemoja, jolloin teemat solussa on merkittynä vain ja ainoastaan yksi teema. Tämä tehtiin analyysin laskennallisen käsittelyn helpottamiseksi. Taulukossa 4.2 on esimerkkinä osa yhden keskustelun taulukoinnista.

### 4.3.2 Python-skriptit dialogirakenteen ja roolisummien visualisointiin

Kaikki keskustelut tallennettiin yhteen taulukkolaskentatiedostoon, josta ne tallennettiin 16 erilliseen .csv-tiedostoon. Näitä tiedostoja käytettiin teemojen yhteismäärien laske-

miseen, keskusteluiden visualisointiin temporaalisena asynkronisena verkkona ja roolien laskemiseen sekä niiden temporaalisen kehityksen visualisointiin.

Python-skripti kävi kaikki csv-tiedostot läpi ja luki ne yksikerrallaan DataFrame-tietueeksi, jossa oli samat rivit ja sarakkeet kuin alkuperäisessä tiedostossa. Tämän jälkeen DataFrameasta poistettiin saman viestin useat rivit, sillä teemoilla ei ollut tässä vaiheessa väliä ja haluttiin tarkastella vain viestien välisiä yhteyksiä. Tämän jälkeen skripti jakaantui kolmeen osaan, joista ensimmäinen teki tex-tiedoston joka sisälsi  $\text{\LaTeX}$ -koodia tikz-kuvaa varten. Toinen laski roolien lukumäärät sopivalla aika-askeleella ja kolmas osa piirsi matplotlib-kirjaston avulla matriisi esityksen roolisummista.

Tikz-kuvan piirtämiseksi luotiin ensin tyhjä matriisi, jonka ensimmäisellä rivillä oli keskustelijoiden nimimerkit ja pystyakselilla aika-asteikko. Skripti luki jokaisen viestin lähettäjän ja lähetysajan, joiden avulla viesti merkattiin harmaalla pallolla oikealle kohdalle. Lisäksi jos viesti oli vastaus toiseen viestiin, piirrettiin nuoli pallosta toiseen. Nuolien väri määräytyi episodin mukaan.

Roolien laskemista varten luotiin ensin viereisyysmatriisi nollamatriisina jokaiselle keskustelun episodille. Tämän jälkeen jokainen vastaus luettiin DataFrameasta episodeittain, ja jos viesti oli vastaus toiselle keskustelijalle, sijoitettiin viereisyysmatriisiin vastaavan yhteyden kohdalle 1. Viereisyysmatriisista laskettiin roolien 1-9 lukumäärät teoriaosuudessa esitettyjen laskukaavojen avulla 1-3 vastauksen välein. Nämä niin sanotut roolisummat sijoitettiin omaan taulukkoon. Lisäksi vastaavat roolisummat laskettiin ja tallennettiin omaan taulukkoon viereisyysmatriisin käänteismatriiseille, jotta voitiin tarkastella roolisummia myös tiedon liikkumissuunnassa.

Roolisummien taulukosta laskettiin normitetut roolisummat jakamalla ko. taulukon jokainen rivi vaakavektorilla, jossa oli neljän solmun täydellistä verkkoa vastaavat roolisummat. Nämä tulokset tallennettiin omaan taulukkoon.

Skriptin kolmas osuus käytti edellä tuotettuja taulukoita roolisummien matriisiesitysten piirtämiseen ja tallentamiseen. Tähän käytettiin matplotlib-kirjastoa ja ns. lämpökartan tyylistä esitystä. Lämpökartssa kuvan väri määräytyy sitä vastaavassa solussa olevan lukuarvon mukaan. Skripti siis luki jokaisen taulukon ja tuotti niistä kuvan, joka tallenneettiin png-tiedostoksi. Kuvan vaaka-akselilla on roolit yhdestä yhdeksään ja pystyakselilla on vastauksien lukumäärä, joka siis käytännössä vastaa keskustelun ajallista etenemistä. Kaikkiin kuviin käytettiin samaa väriskaalaa, jossa nollaa vastasi valkoinen ja musta vastaa suurinta mahdollista arvoa eli roolisummissa sataa roolia ja normitetuissa tuloksissa yhtä eli maksimimäärää rooleja. Luettavuuden parantamiseksi skaalan välillä olevat arvot saavat sinisen eri sävyjä.

Erillinen skripti kirjoitettiin teemojen laskemiseksi. Yksinkertainen skripti avasi vuorollaan jokaisen csv-tiedoston ja laski jokaisesta keskustelusta episodikohtaiset teemat. Sama skripti laski lisäksi yhteen kaikki niissä esiintyneet teemat. Näin saatiin selville

teemojen lukumäärät, joiden avulla pystyttiin piirtämään pylväsdiagrammit episodeissa esiintyneistä episodeista sekä yhteensä kaikista episodeista.

Työn alkuvaiheilla oli suunnitelmassa tarkastella laajempaa aineistoa tai ainakin taulukoida keskustelut Moodle-alustalta valmiiksi. Käsini tehtynä tämä olisi vienyt kohtuuttomasti aikaa, mutta Moodle-alustasta on mahdollista tehdä kurssikohtainen varmuuskopio. Varmuuskopio pitää sisällään myös keskustelut xml-tiedostoissa.

Tämän pohjalta kirjoitettiin python-skripti joka käyttää beautiful soup -pakettia, jolla keskustelut voitiin taulukoida. Skripti lukee kerrallaan yhden xml-tiedoston ja tarkistaa siellä olevia tietoja viestien lähettäjistä, lähetysajoista ja sisällöistä. Viestit sijoitetaan vastaavan tyyppiseen taulukkoon, joka edellä oli tehty käsini, mutta nyt teemojen sijaan niissä on viestien sisällöt. Lopuksi taulukko tallennettiin csv-tiedostoksi.

## 5. Tulokset

Keskusteluissa oli yhteensä 638 viestiä, joista 535 oli vastauksia. Ne olivat jakaantuneet 35:een erilliseen keskusteluepisodeihin. Taulukossa, 5.1, on esitetty keskusteluepisodeissa esiintyneiden vuorovaikutusten lukumäärät.

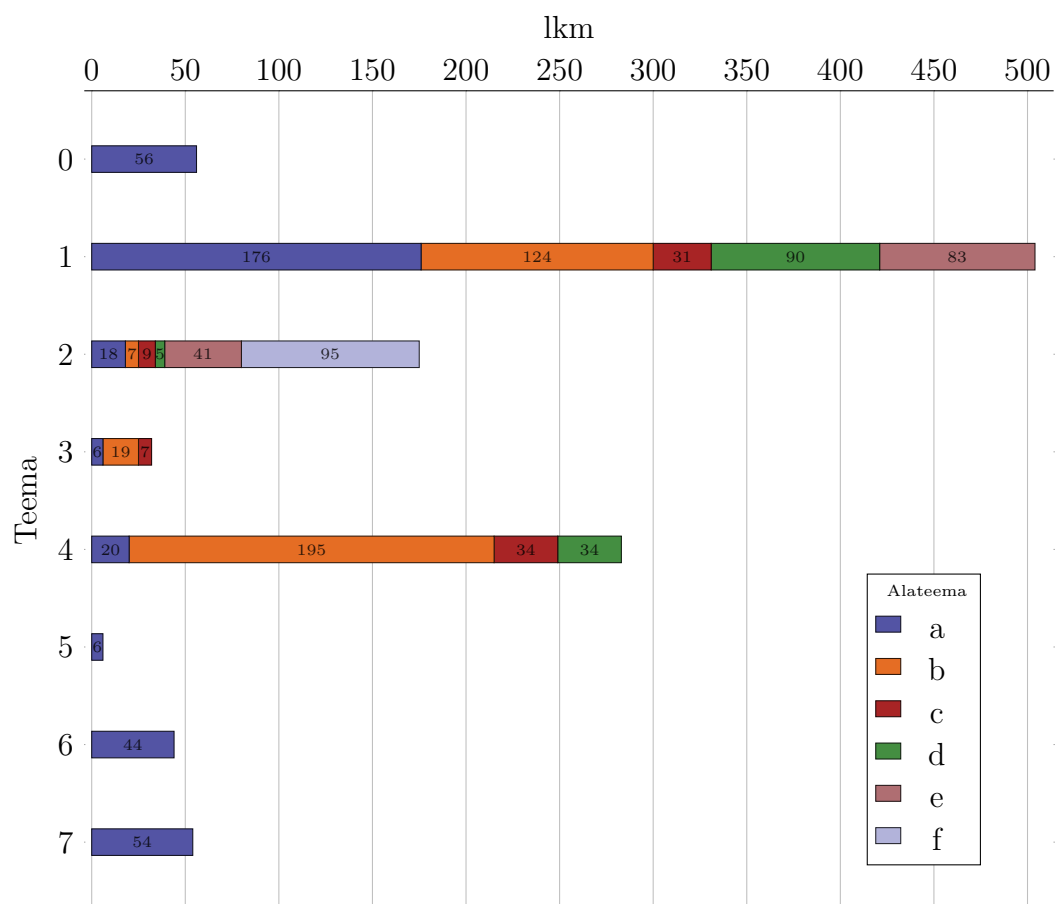
**Taulukko 5.1:** Ryhmien vuorovaikutusten lukumäärät keskusteluepisodeittain. Episodit on erotettu pilkulla toisistaan.

Harj.	A	B	C	D
1	12, 22	23	19	30
2	18, 11	22, 7	26, 12	33, 20
3	17, 10, 11	17, 7	15	27, 15
4	28	12,12	6	27, 8

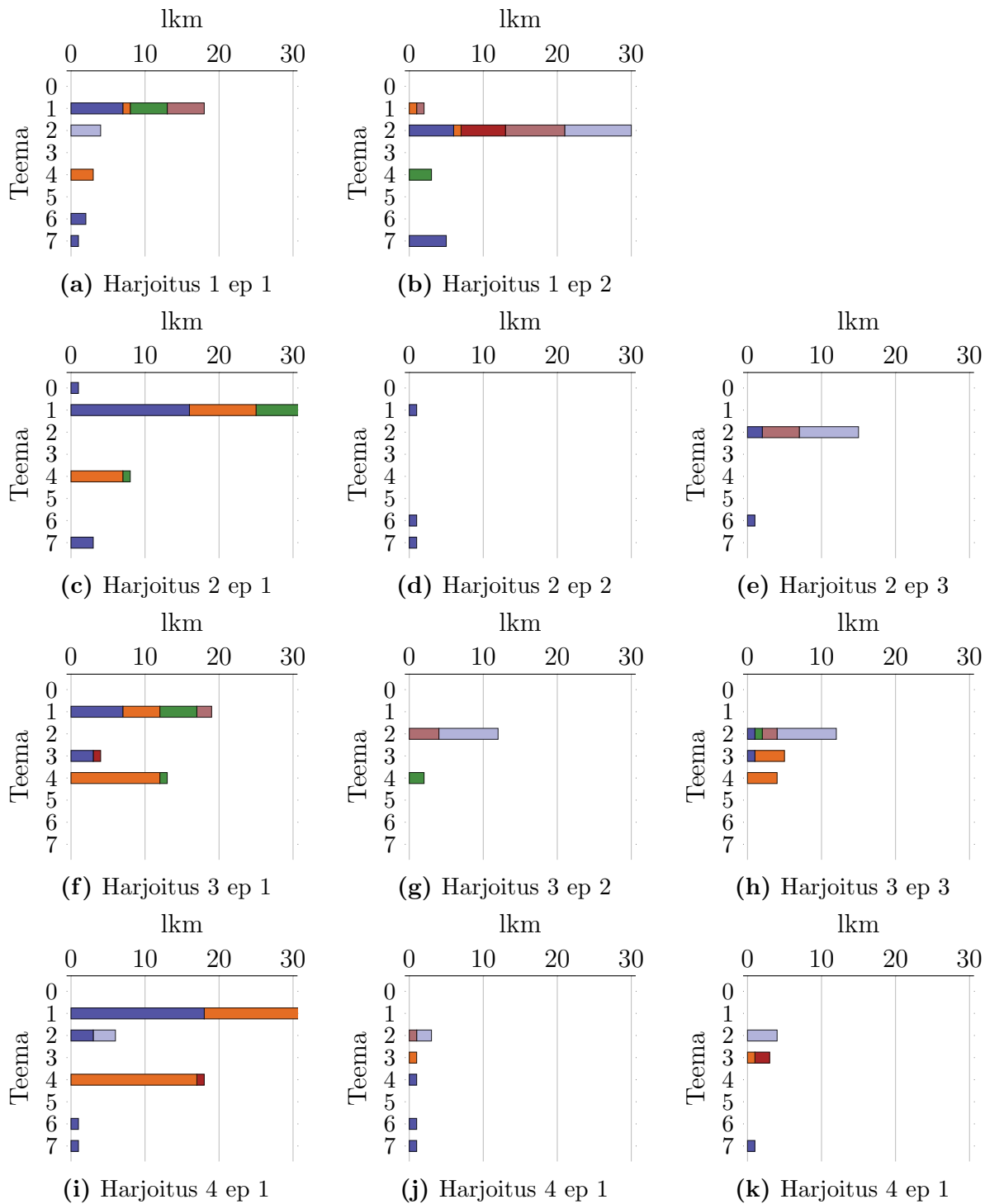
### 5.1 Teemat ja episodit

Teemojen lukumäärät on esitetty pylväsdiagrammissa kuvassa 5.1. Alateemat ovat eritelty eri väreillä pääteeman sisällä. Kuvissa 5.2-5.5 on myös esitetty jokaisessa keskusteluepisodissa esiintyneet teemat. Näissä kuvissa on alateemat merkitty samalla tavalla kuin kuvassa 5.1.

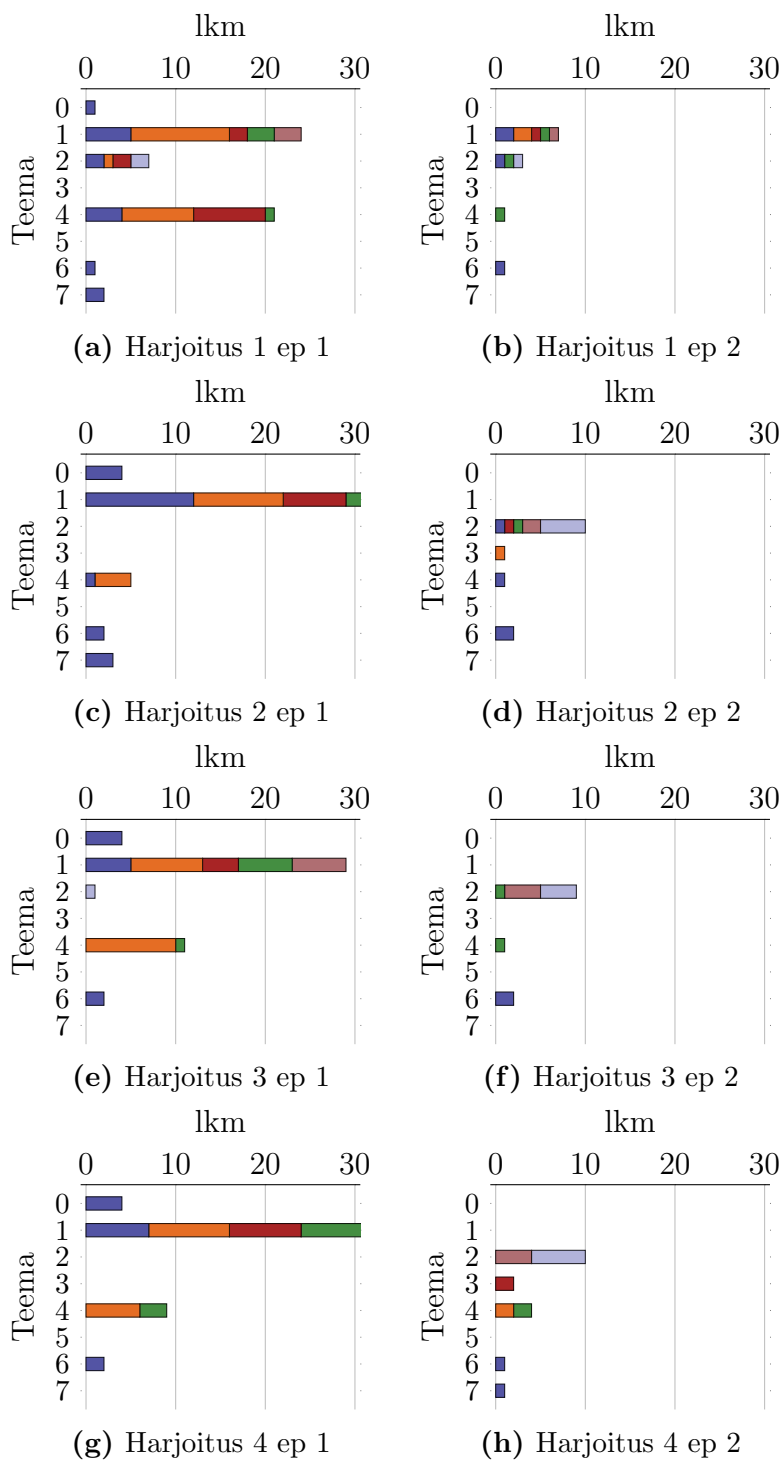




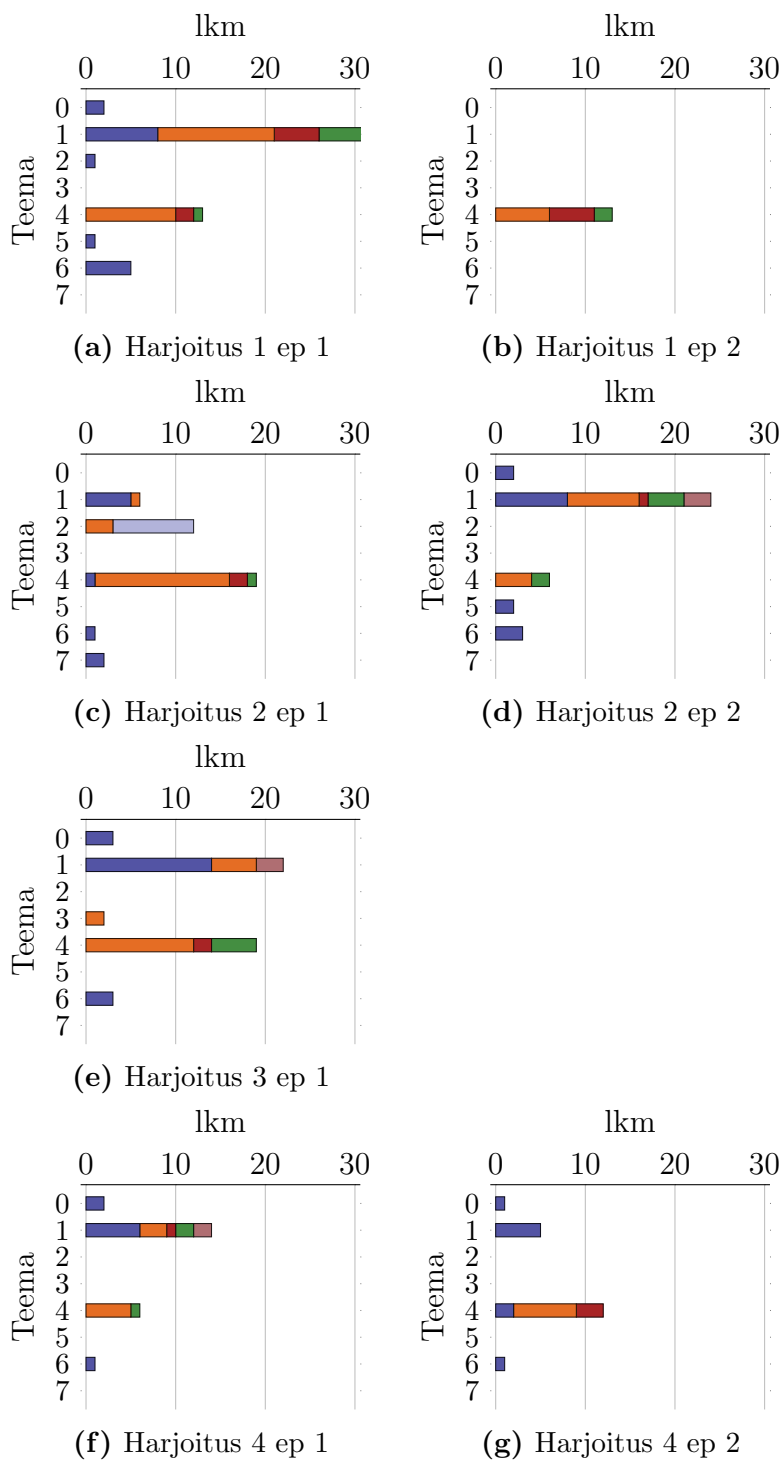
**Kuva 5.1:** Pylväsdiagrammi teemojen jakautumisesta keskusteluissa. Alateemat on merkitty erivärisinä osuuksina pääteeman sisällä.



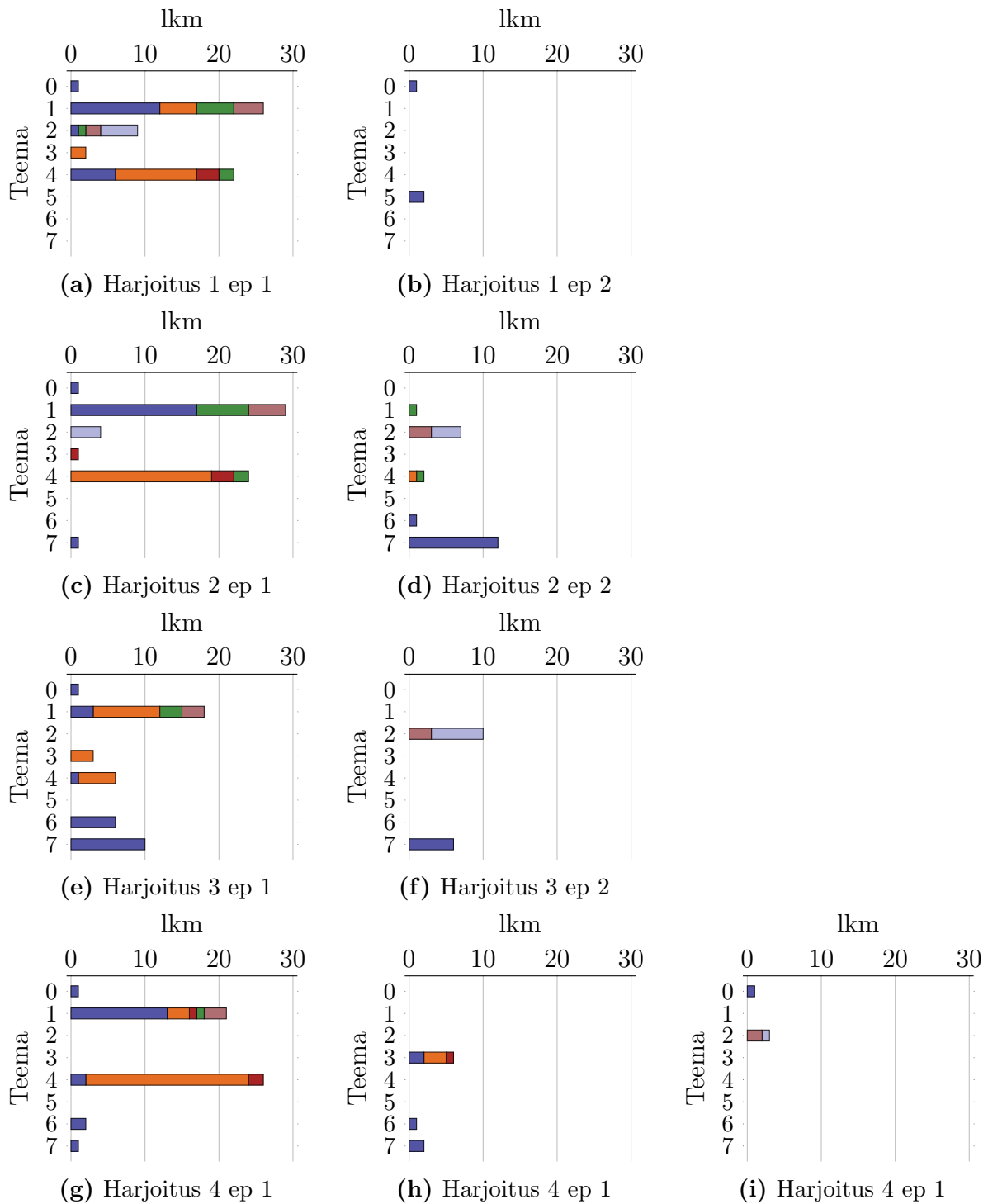
Kuva 5.2: Ryhmä A



Kuva 5.3: Ryhmä B



Kuva 5.4: Ryhmä C



Kuva 5.5: Ryhmä D

## 5.2 Asynkroniset temporaaliset verkostot

Verkkokeskusteluiden dialogirakenteet esitetään asynkronisena temporaalisena verkostona. Ne ovat esitetty kuvissa 5.6-5.9 keskusteluittain siten, että jokaisessa kuvassa on aina yhden ryhmän keskustelut tehtäviin 1-4 liittyen. Yhdessä keskustelussa on aina ensimmäisellä vaakarivillä keskusteluun osallistuneet henkilöt ja pystyakselilla on aika ordinaaliasteikolla.

Keskusteluissa viesti on merkitty harmaalla palolla ja keskustelun eteneminen on esitetty nuolilla. Nuolten väri kuvaa, mihin episodiin ne liittyvät.

Tarkastellaan tarkemmin ryhmän A keskustelua harjoituksesta kaksi, joka on kuvassa 5.6 toisena. Tätä on helppo seurata, sillä A4 ei käytännössä osallistu koko keskusteluun. Ensimmäiseen keskusteluepisodeihin osallistuu siis jäsenet A1-A3 ja heidän välillään on yhteensä 18 vuorovaikutusta, kun ei huomioida viestejä itselle, aikavälillä 3-28.

Taulukossa 5.2 on esitetty rakenteelliset roolityypit jokaiselle jäsenelle sitä mukaan, kun ne syntyvät siten, että roolin katsotaan syntyneen vastauksen ajanhetkellä. Ensimmäinen vastaus on tehty ajanhetkellä viisi ja se on A2:lta A1:lle, jolloin A1:n roolityyppi on yksi eli nielu. Vastaavasti A2 on tällöin lähde eli sen on roolityypiltään kaksi. Toisin sanoen viesti ja vastaus muodostavat yhden yksisuuntaisen linkin. A3 ei ole vielä osallistunut keskusteluun, joten hänellä ei ole roolia.

**Taulukko 5.2:** Ryhmän A jäsenten rakenteellisten roolien aikakehitys harjoituksen kaksi keskusteluepisodissa yksi.

t	A1	A2	A3
5	1	2	-
6	1	5	1
8	4	5	6
9	4	4	6
10	4	8	8
11	7	9	8
12	9	9	9

Seuraavalla ajanhetkellä A2 vastaa A3:lle, jolloin A3:n roolityyppi on nielu. A1:n rooli ei nyt muutu, mutta A2 rooli vaihtuu roolityypiltään viideksi eli 2-lähteeksi. Edelleen kuitenkin A2:lla on funktionaalisena roolina myös tyyppi kaksi. Kolmannessa vuorovaikutuksessa, kun A3 vastaa A1:lle, muodostuu A3:n rooliksi tyyppin kuusi mukainen broker tai relay- eli välittäjätyyppinen. Toisin sanoen A3:n kautta muodostuu välillinen yhteys A1:n ja A2:n välille.

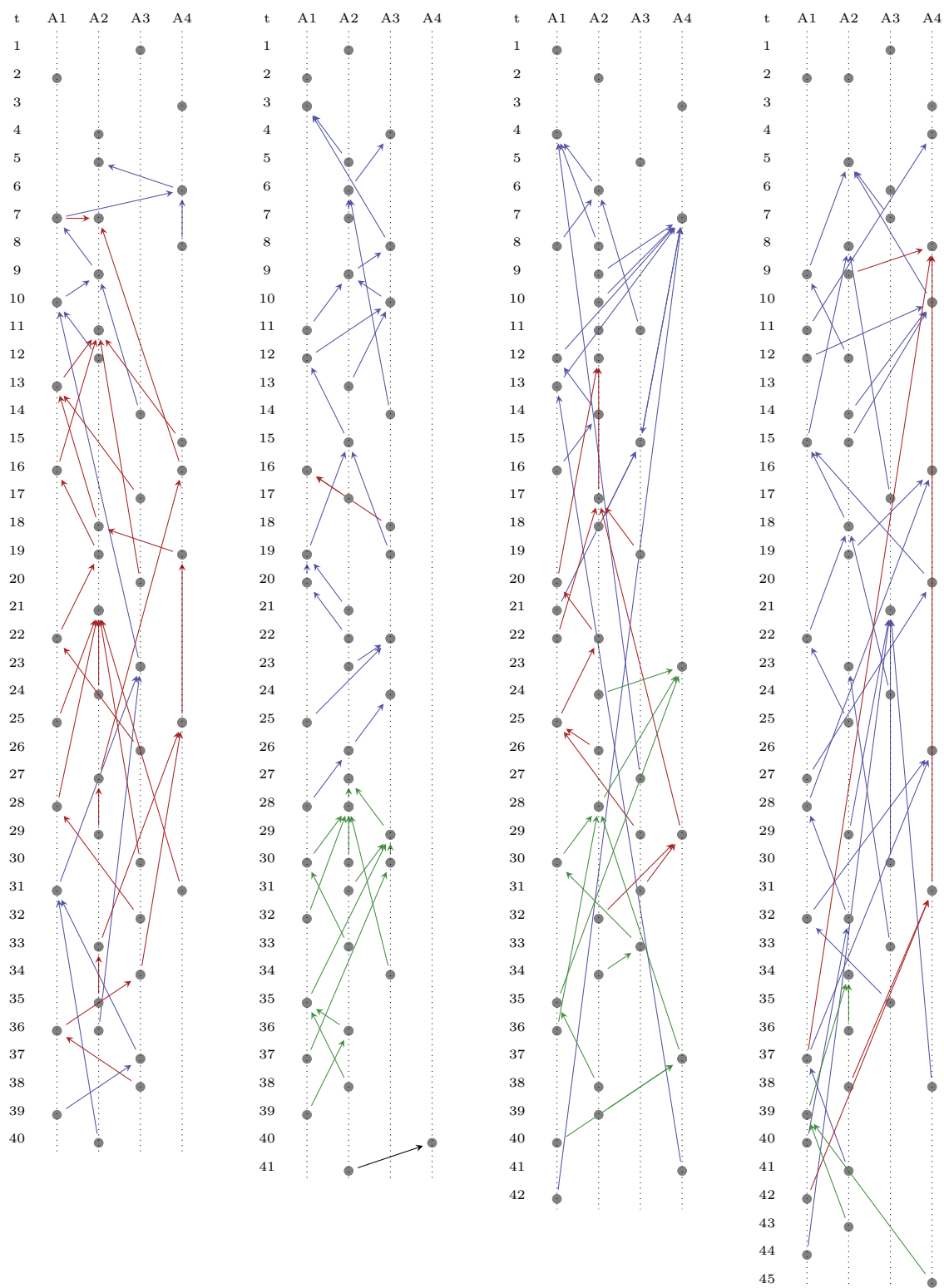
Taulukosta käy myös ilmi, että ajanhetkellä yhdeksän ei tapahdu muutoksia rooleis-

sa. Tämä johtuu siitä, että A2 on jo aiemmin vastannut A3:lle, joten verkon rakenteessa ei tapahdu muutoksia.

Heti tämän jälkeen kuitenkin A3 vastaa takaisin A2:lle, jolloin heidän välilleen syntyy vastavuoroinen yhteys, joka näkyy myös heidän roolissa. Nyt heidän molempien roolityyppi on 8 eli heillä on keskenään vastavuoroinen suhde ja molemmat toimivat myös lähteenä A1:lle.

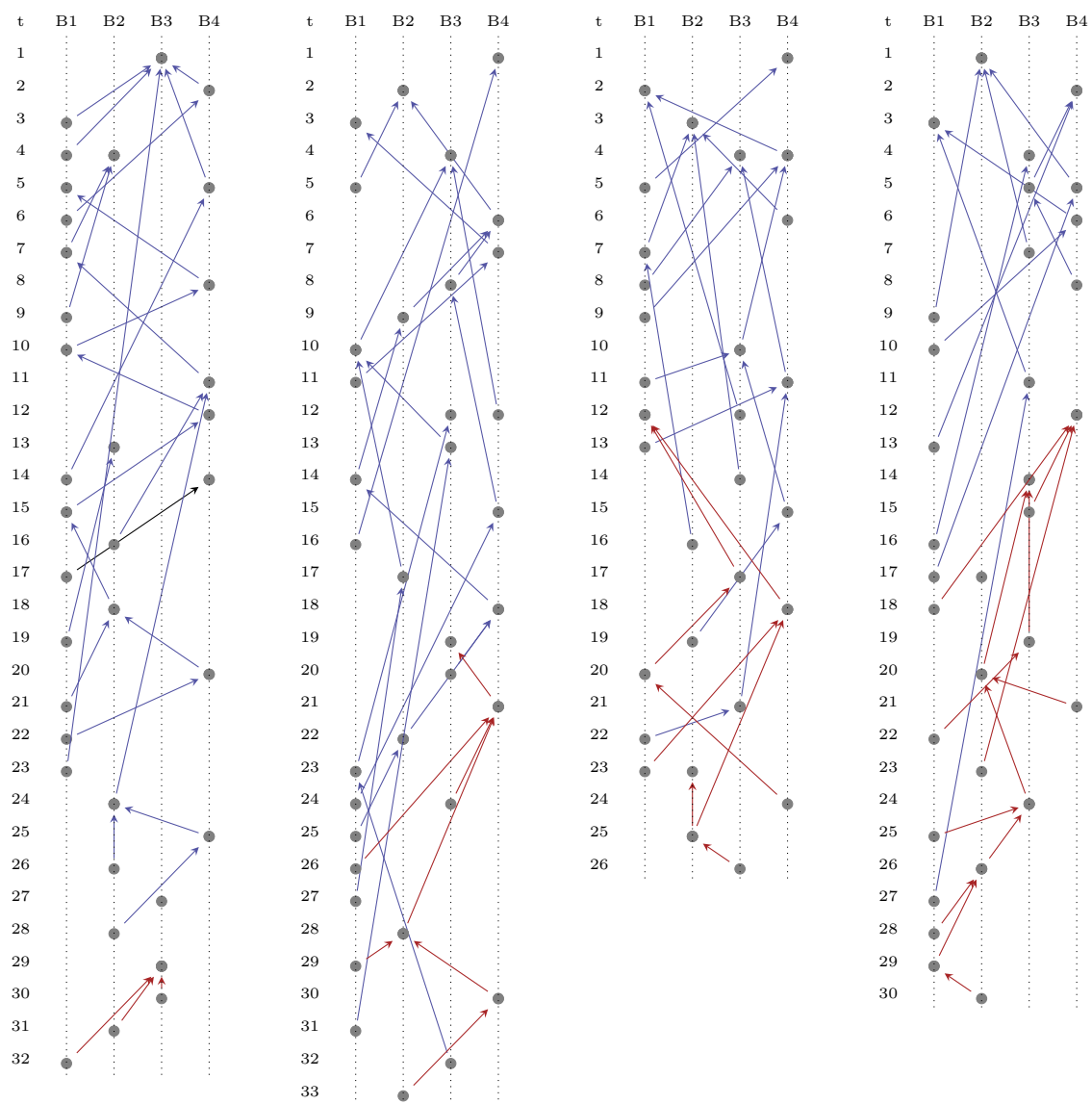
Taulukko etenee vastaavasti, mutta loppuu jo ajanhetkeen 12 eli seitsemänteen vuorovaikutukseen. Tämä johtuu siitä, että tässä vaiheessa kaikkien triadin jäsenten roolityyppi on yhdeksän, joten heidän välinen verkko on täysi. Toisin sanoen verkon rakenne ei voi enää muuttua ja siten lopuilla vuorovaikutuksilla ei ole enää merkitystä roolisummissa. Tämä nähdään myös suoraan kuvassa 5.10f, jossa on kyseisen keskusteluepisodin roolisummat normitettuna. Siinä roolisummaa kuvaava väri ei enää muutu seitsemännen vuorovaikutuksen jälkeen.

Vastaava tulkinta voitaisiin tehdä mille tahansa keskustelulle, mutta se olisi työlästä, etenkin kun kaikki neljä jäsentä osallistuvat keskustelun. Esimerkiksi ryhmän D keskustelun kaksi episodissa kaksi kaikkien neljän osallistujan välille muodostuu yhteys heti neljän vuorovaikutuksen jälkeen. D1 on jokaisessa kolmessa triadissaan roolissa 2-nielu ja muut jäsenet vastaavasti lähteitä. Neljännessä triadissa olisi linkki vain D2:sta D4:ään. Manuaalista analyysiä voitaisiin jatkaa, jolloin päästäisiin tulokseen, että ajanhetkellä 45 verkko olisi täysi. Tällöin jokaisen jäsenen välillä on vastavuoroinen suhde, joka taas nähdään normitetusta roolisummista kuvassa 5.13f.

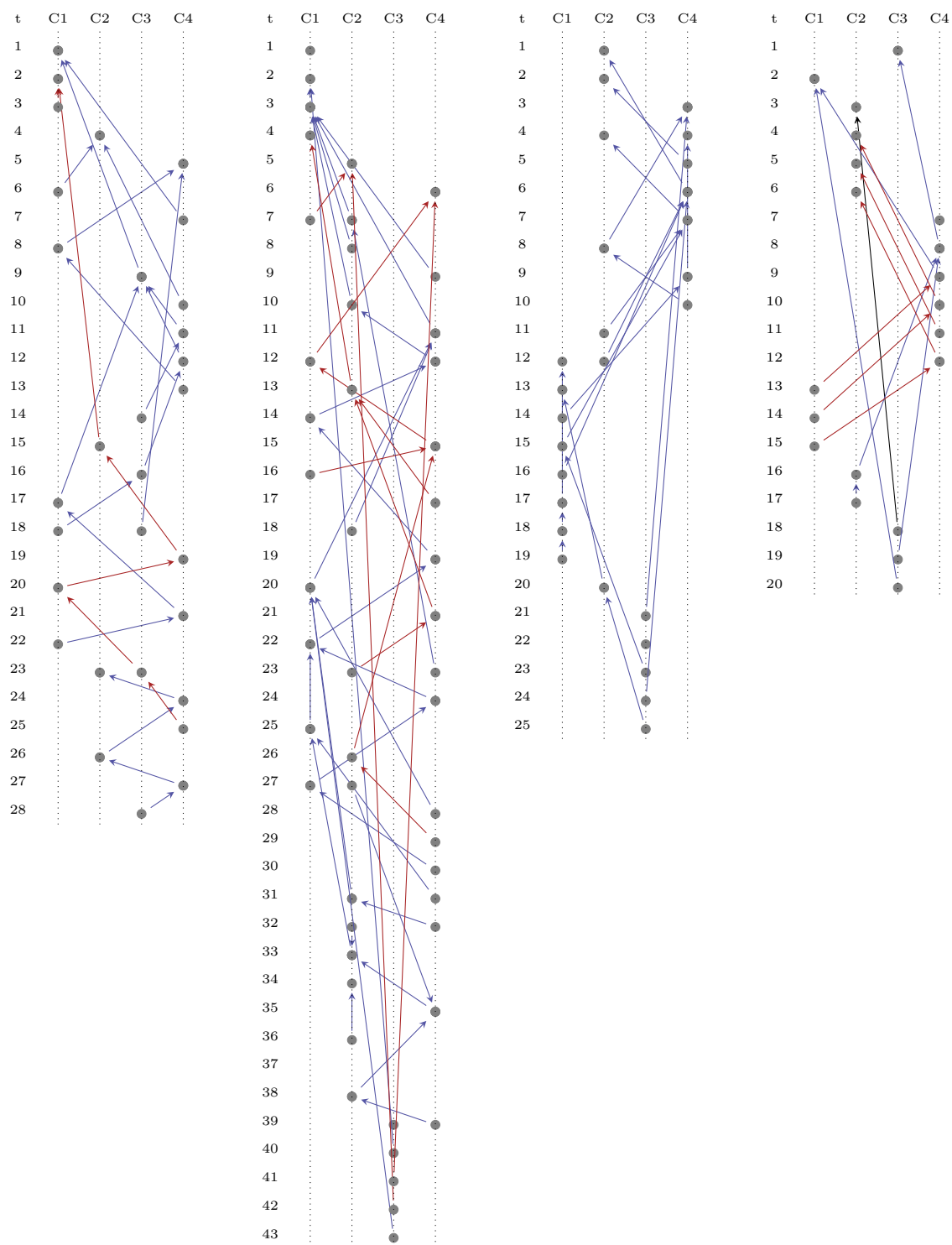


Kuva 5.6: Ryhmän A keskustelut harjoituksista 1-4.

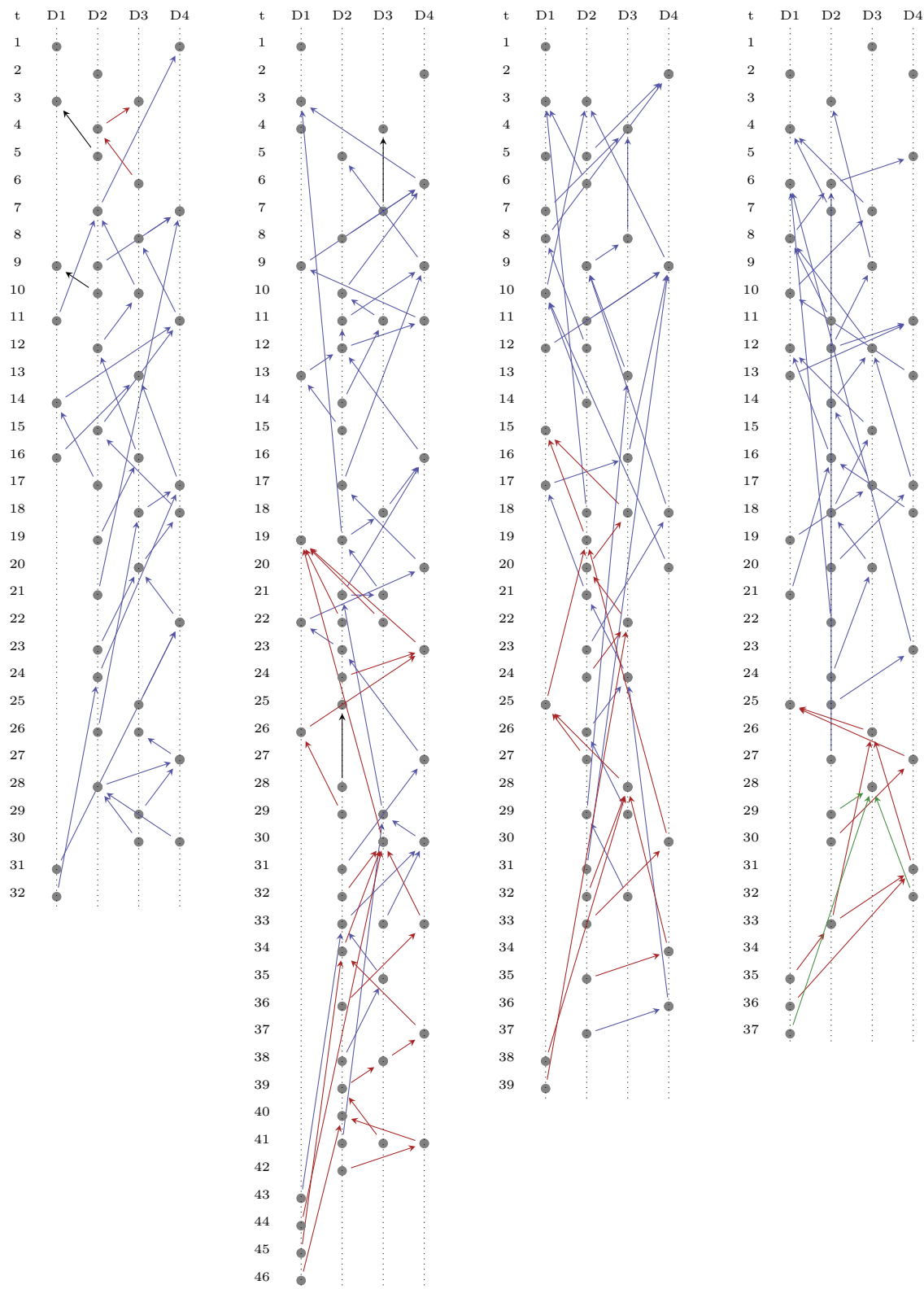




**Kuva 5.7:** Ryhmän B keskustelut harjoituksista 1-4.



**Kuva 5.8:** Ryhmän C keskustelut harjoituksista 1-4.



Kuva 5.9: Ryhmän D keskustelut harjoituksista 1-4.

## 5.3 Roolit

Edellisessä kappaleessa esitettyjen syiden takia roolisummat on esitetty matriisiesityksinä. Edellinen kuitenkin osoittaa kuinka dialogirakenteen rakennetta ei täysin huomioida muodostettaessa viereisyysmatriisia roolisummien laskemista varten.

Koska tarkoituksena on vertailla sosiaalisissa verkoissa esiintyviin rakenteisiin esitetään keskusteluissa esiintyneet roolisummat viestien mukaisessa suunnassa. Kuvissa 5.10-5.13 on ryhmäkohtaisten keskusteluiden roolisummat episodeittain. Jokaisesta episodista on ensin esitetty koko summat ja toisena normitetut summat.

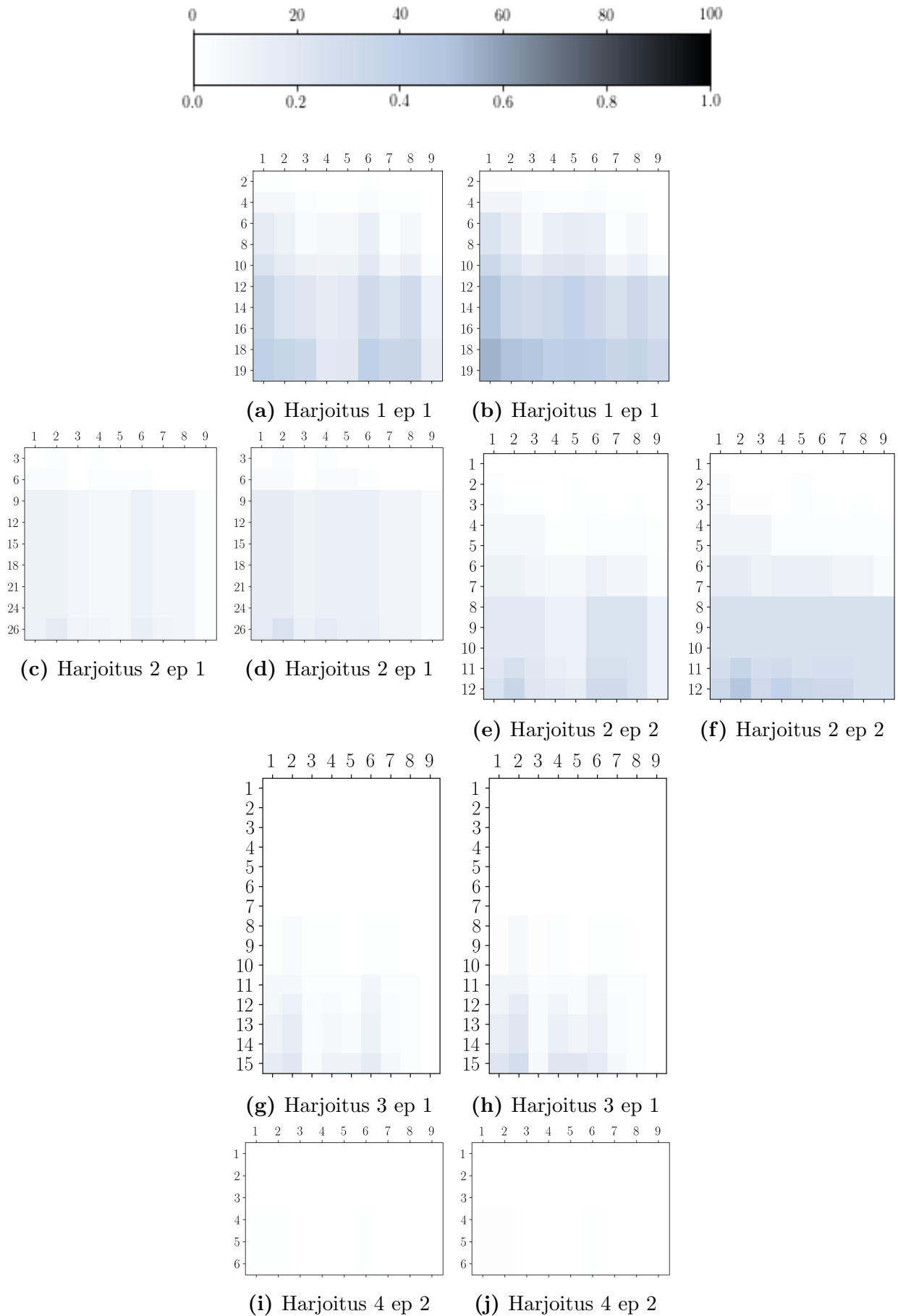
Kuvissa vaaka-akselilla on roolit 1-9 ja pystyakselilla niiden lukumäärän aikakehitys. Kuvissa lukumäärät on esitetty lämpökarttana eli mitä tummempi solu on sitä suurempi roolin lukumäärä keskustelussa on ollut. Rooleina lukumääräväli on 0-100 ja normittuna 0-1, koska normitetut tulokset kuvaavat käytännössä osuutta maksimista. Aika-askel kuvaa siihen asti esiintyneitä vuorovaikutuksia eikä siten ole suoraan verrannollinen asynkronisten verkkojen aika-askeliin.



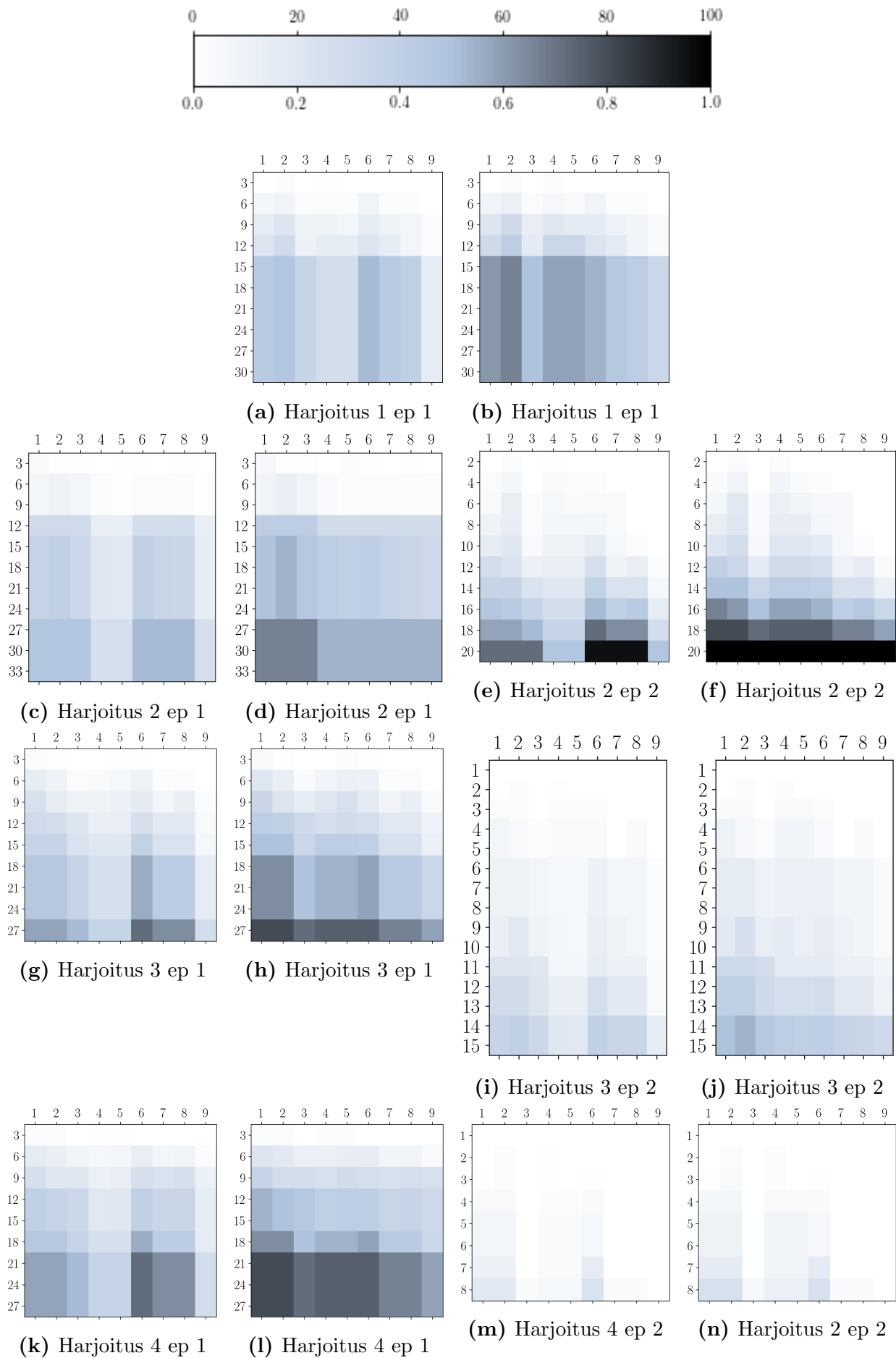
**Kuva 5.10:** Ryhmän A roolisummat episodeittain absoluuttiset vasemmalla ja normitettuna oikealla. Vaaka-akselilla on roolityypit ja pystyakselilla on vuorovaikutusten lukumäärä episodissa.



**Kuva 5.11:** Ryhmän B roolisummat episodeittain absoluuttiset vasemmalla ja normitettuna oikealla. Vaaka-akselilla on roolityypit ja pystyakselilla on vuorovaikutusten lukumäärä episodissa.



**Kuva 5.12:** Ryhmän C roolisuunnat episodeittain absoluuttiset vasemmalla ja normitettuna oikealla. Vaaka-akselilla on roolityypit ja pystyakselilla on vuorovaikutusten lukumäärä episodissa.



**Kuva 5.13:** Ryhmän D roolisummat episodeittain absoluuttiset vasemmalla ja normitettuna oikealla. Vaaka-akselilla on roolityypit ja pystyakselilla on vuorovaikutusten lukumäärä episodissa.



## 6. Pohdinta

Tämän työn tavoitteena oli löytää työkalut verkkokeskustelun auki purkamiseen ja analysoimiseen. Tätä varten tehtiin teemattinen analyysi, jotta saataisiin kuva mistä on keskusteltu. Keskustelun rakenteesta taulukoitiin vuorovaikutukset, jotta rakenne pystyttiin esittämään graafisesti asynkronisena temporaalisena verkkona. Lisäksi sovellettiin sosiaalista verkostanalyysiä keskusteluissa olevien roolien esittämiseksi.

Koska tehtiin temaattinen analyysi, ei tuloksia voida sellaisenaan tarkastella kvantitatiivisesti. On myös oleellista huomata, että yhdessä viestissä saattaa olla yksi tai useampi teema, eikä viestien pituuteen ole otettu ollenkaan kantaa. Nämä tulisi ottaa jollain tapaa huomioon, jos haluttaisiin tehdä tarkempi kvantitatiivinen analyysi viestien sisällöistä.

Jos viestien sisältöjä halutaan tutkia tarkemmin, olisi kenties suotavaa käyttää tarkkaa sisällönanalyysiskeemaa temaattisen analyysin sijaan. Tässä työssä käytetyn temaattisen analyysin tuloksia ja esitettyjä teemoja voidaan kuitenkin käyttää tulevien analyysien pohjana.

Teemojen tehtävä oli lähinnä saattaa viestit episodeiksi ja varmistaa laadullisesti, mitä asioita keskustelut pitivät sisällään. Jos keskustelut eivät olisi olleet tehtävänannon mukaisi, ei olisi ollut mielekästä jatkaa analyysiä sosiaalisiin vuorovaikutuksiin. Tavoitteena oli kuitenkin löytää työkalut nimenomaan verkko-oppimisen sosiaalisiin ulottuvuuksiin, joten jos keskustelut eivät liittyisi tehtävään, ei myöskään sosiaaliset vuorovaikutukset koskisi oppimista.

Episodeissa haasteena olivat useat keskustelun aloitukset. Myöhemmillä kursseilla keskustelun ohjeet ovat olleet strukturoidumpia, jolloin episodit todennäköisesti nähtäisiin suoraan aloitetuista keskusteluista. Alustava keskusteluiden tarkastelu näyttäisi viittaavan tähän suuntaan.

Teemat ovat melko lailla episodikohtaisia, mutta niin pitääkin olla, koska keskusteluepisodit määriteltiin temaattisesti eheiksi kokonaisuuksiksi. Toki tulokset osoittavat, että ns. ensisijaisen teeman lisäksi keskusteluissa esiintyvät rinnalla myös muuta keskustelun teemat. Lisäksi nähdään, että keskustelu on ollut pääasiassa tehtävänannon mukaista ja vain vähän muunlaista keskustelua heijastelevia teemoja esiintyy keskusteluissa, joka on yhtenevää kirjallisuuden kanssa (Schellens & Valcke, 2006).

Selkeästi kuitenkin pääpaino keskustelussa on ollut enemmän kaavioiden vertailua ja

vain toissijaisesti opetukseen ja didaktiseen rekonstruktion liittyvää. Tämä saattaa mahdollisesti johtua ohjeistuksen melko vapaasta muodosta ja aikarajoitteista, jolloin opiskelijat ovat ehtineet keskustella vain ohjeistuksen alun mukaisesti. Huomionarvoista on, että teema 4c nousee esiin teemojen yksi ja kaksi lisäksi. Teema 4c koskee tehtävän haasteellisuutta ja opiskelijat ilmeisesti pitivät tehtävien tekemistä englanninkielisten artikkeleiden pohjalta haastavana.

Teemojen ja verkkokeskustelun puurakenteen avulla saatiin keskustelun dialogirakenne taulukoitua. Taulukon avulla dialogirakenne onnistuttiin esittämään asynkronisena temporaalisena verkkona, jonka avulla dialogirakenne voidaan helposti hahmottaa yhdellä silmäyksellä. Kuvien avulla on mahdollista arvioida keskustelun laatua sekä havaita esiintykö keskustelussa asynkronisuutta.

Esimerkiksi on selkeästi nähtävillä, että ryhmän C kaksi viimeisintä keskustelua ovat olleet jokseenkin vajavaista ja yksinkertaista. Vastaavasti taas ryhmän D kaikki keskustelut ovat olleet kohtalaisen pitkiä ja monimuotoisia.

Keskustelut täyttävät automaattisesti asynkronisuuden ehdoista kaksi: missä vain ja suora viittaus edelliseen viestiin. Kurssijärjestelyn osalta keskustelu ei kuitenkaan ole luonteeltaan milloin vain, sillä opiskelijoita ohjeistettiin olemaan kirjautuneena keskustelualueelle samaan aikaan kahden tunnin ajan.

Keskustelussa esiintyy kuitenkin asynkronisuutta, jossa viitataan selkeästi aiempaan viestiin. Erityisen selkeä esimerkki asynkronisuudesta on ryhmän B harjoituksen yksi keskustelussa. Opiskelija B2 vastaa opiskelijalle B4 samaan viestiin ( $t=11$ ) kahdesti ( $t=16$ ,  $24$ ). B2 aloittaakin toisen viestinsä ”*Jäin miettimään tuolta aiemmasta kommentista...*” eli hän palaa uudelleen B4:n esittämään asiaan.

Tämä antaa osviittaa, siitä että mahdollisesti opiskelija on muun keskustelun ohessa ehtinyt reflektoida ko. viestin sisältöä ja tämän jälkeen hän palaa aiheeseen. Voisi olla mielenkiintoista tarkastella onko keskusteluissa laadullista eroa ja palataanko edellisen aiempiin viesteihin, jos keskustelua käytäisiin vaikkapa viikon aikana tyyliä missä ja milloin vain. Ongelmaksi saattaa kuitenkin muodostua keskustelun ohjeistaminen ja keskustelun määrän varmistaminen ilman, että opiskelijat lisäävät alueelle vaaditun määrän viestejä heti aluksi.

Roolien määrää ja ajallista kehittymistä kuvaavat matriisiesitykset tuovat selkeästi esille funktionaalisten roolein jakauman ja niiden ajallisen kehityksen. Yleisesti ottaen keskusteluissa esiintyneiden roolien lukumäärä kasvaa keskustelun pituuden myötä. Tämä toki on täysin triviaalia ja esimerkiksi edellä mainituissa ryhmän C keskusteluissa ei käytännössä esiinny rooleja juurikin lyhyiden ja yksinkertaisten keskusteluiden seurauksena.

Toisaalta ryhmän D keskustelun kaksi toinen episodi on ainoa, jossa esiintyvät kaikki mahdolliset roolit eli sen verkko on täydellinen. Vaikuttaisi siis siltä, että keskustelun dialogirakenteen graafisen esityksien havaintojen ja roolien lukumäärillä on yhteys. Tä-

mä tarkoittanee lähinnä sitä, että molempia esityksiä voidaan käyttää keskustelun laadun pikaiseen arviointiin vaikkakin ne kuvaavat hieman eri asioita. Graafisella kuvauksella saadaan nopea yleissilmäys keskusteluun ja roolien avulla nähdään helposti, kuinka vastavuoroista keskustelua on ollut ryhmän sisällä.

Voisikin olla mielenkiintoista vertailla keskustelun rakennetta ja esiintyneitä rooleja oppimistuloksiin. Kurssilla, jolta käytetty aineisto on kerätty, ei keskustelun näennäisellä laadulla tai vastavuoroisuuden määrällä ollut merkittävää korrelaatiota tehtävistä saatuihin pisteisiin. Yhdestä kurssista saatava aineisto on kuitenkin liian pieni, jotta tästä voitaisiin tehdä minkäänlaisia johtopäätöksiä. Lisäksi, jos tätä haluttaisiin tutkia tarkemmin, pitäisi myös analysoida argumenttikaavioiden tai didaktisten rekonstruktioden sisältöjä. Nyt kun keskustelun rakenteet ja roolit saadaan helposti ja nopeasti selville, olisi helppo tarkastella niiden yhteyksiä oppimistuloksiin myös usean kurssin tulosten pohjalta.

Roolit eivät ole täysin suoraan verrattavissa malleihin sosiaalisista rakenteista, sillä malleissa keskitytään usein kokonaisiin triadeihin tai motiffeihin eikä niinkään yksittäisiin rooleihin. Esitetyillä roolisummien kuvilla voidaan kuitenkin esimerkiksi tarkastella esiintyykö keskustelussa resiprokaalisia roolityyppejä 3 ja 7-9. Lisäksi tämän tyyppisellä roolien esittämällä saadaan helposti esille niiden temporaalinen kehitys.

Vain yhdeksässä keskusteluepisodissa vuorovaikutuksia on 22 tai useampi, joten Skvoretzin vaatimus vuorovaikutusten lukumäärästä hierarkisen statusjärjestyksen syntymiselle ei täyty kaikissa keskusteluissa. Lyhyissä keskusteluissa pääpaino onkin yksisuuntaisissa suhteissa eli niissä esiintyy tyyppien yksi, kaksi ja kuusi mukaisia rooleja. Eniten esiintyy tyyppiä kaksi eli viestin lähettäjiä, joka sinänsä on triviaalia, tosin johtaja roolit, tyyppi 4, eivät merkittävästi nouse esiin vaikkakin esiintyvät hieman aiemmin.

Lyhyissä keskusteluissa esiintyy siis syklisiä rooleja eli tyyppin kuusi mukaisia kaverin kaveri -rooleja, joiden tulisi kehittyä kohti resiprokaalisia rooleja. Käytetty laskentamalli ei kuitenkaan mahdollista tyyppien 7-9 rooleja ilman, että samalla ei esiintyisi tyyppiä kuusi. Tämä johtuu siitä, että tarkasteltiin funktionaalisia rooleja, jolloin tyyppi kuusi jää ikään kuin taustalle muiden tyyppien muodostuessa. Voidaan siis olettaa, että käytännössä mahdolliset suhteet ovat vastavuoroisia eivätkä enää syklisiä siinä vaiheessa kun resiprokaaliset roolit ilmenevät.

Pitkissä keskusteluissa kuitenkin esiin nousevat vastavuoroiset suhteet oletettujen mallien mukaisesti. Tosin pitkä keskustelu ei kuitenkaan suoraan takaa, että kaikki suhteet olisivat resiprokaalisia. Esimerkiksi ryhmän D harjoituksen kaksi ensimmäisessä episodissa on 33 vuorovaikutusta, mutta rooleja ei esiinny maksimimäärää. Vastaavasti toisessa episodissa on vain 20 vuorovaikutusta ja kaikki suhteet ovat vastavuoroisia. Tämän toki voi selittää se, että ryhmän rakenteisiin on jo vaikuttanut aiempi, ensimmäisessä episodissa käyty keskustelu.

Rooleja ja niiden aikakehitystä voisikin tarkastella koko keskustelusta ilman jakoa

episodeihin tai jopa koko kurssin aikana käytyä keskustelua. Tuolloin myös erottelevuus saattaisi riittää kaikkiin kolmeenkymmeneen rooliin. Toisaalta aiemmat keskustelut eivät näytä vaikuttavat ainakaan positiivisesti myöhempiin keskusteluihin, joista osoituksena ryhmän C suppeat keskustelut harjoituksista kolme ja neljä.

Resiprokaalisuuden lisääntyminen voisi johtua vain teknisistä syistä eli keskustelun pituus lisää todennäköisyyttä viestittelyn menevän ristiin jäsenten välillä. Kuitenkin ryhmän jäsenten kokema hyöty tehtävästä vaikuttaa motivaatoon (E. J. Lawler, 2001). Tosin sanoen, jos tehtävää ei koeta hyödylliseksi, ei juurikaan käydä keskustelua tai keskustelu on suppeaa ja pintapuolista. Tällöin rooleja todennäköisesti esiintyy vähemmän eivätkä ne ole vastavuoroisia. Myös tämä on saattanut olla tekijänä ryhmän C suppeissa keskusteluissa.

Roolien laskeminen on osoittautunut olevan käytettävissä oleva menetelmä roolien analysoimiseen verkkokeskustelussa. Jos jatkossa halutaan tutkia tarkemmin teoreettisten mallien ja roolien yhteyttä, pitää pohtia roolien ja motiifien yhteyttä. Esimerkiksi eräs Skvoretzin stabiili triadi vastaa motiiffia 5, jossa esiintyy roolityyppien 4,5 ja 6 mukaisia rooleja. Pitäisi siis laskea, minkälaisia motiiffeja esiintyy tai sitten selvittää tarkemmin, mitkä roolit ovat teorioiden rakenteissa merkittäviä. Työn tulokset kuitenkin osoittavat, että rooleja laskemalla voidaan löytää yhteys sosiaalisista verkoista löytyviin rakenteisiin.

Käytetty menetelmä tarjoaa pohjaa myös roolien temporaalisen kehityksen tutkimiselle. Tässä työssä ei huomioitu viestien tai vuorovaikutusten ajallista etäisyyttä toisistaan. Tämä voisikin olla toinen mahdollinen tapa lähteä kehittämään roolien analysoimista. Jokaiseen vuorovaikutukseen voitaisiin yhdistää temporaalinen painokerroin, joka mahdollistaisi esimerkiksi asynkronisuuden huomioimisen roolien kehitykseen.

Työn lopullisena johtopäätöksenä voidaan todeta, että onnistuttiin löytämään menetelmät verkkokeskustelun kuvaamiseen ja analysoimiseen. Toisin sanoen työn tavoite täyttyi ja tutkimuskysymyksiin onnistuttiin vastaamaan. Opiskelijat keskustelivat pääasiassa tehtävänannon mukaisista aiheista. Keskustelun rakenne onnistuttiin esittämään graafisesti siten, että sitä voidaan tulkita nopealla silmäilyllä. Vastaavasti myös keskusteluissa esiintyneet roolit ja niiden temporaalinen kehitys saatiin esitettyä graafisesti.

Jatkossa näitä menetelmiä voidaan mahdollisesti käyttää opetuksen arviointiin ja kehittämiseen. Lisäksi kokemukset ja tuotetut python-skriptit luovat pohjaa lisätutkimuksen tekoon. Työssä käytetyt skriptit sekä ohessa syntyneet skriptit Moodle-alustan keskustelun purkamiseen taulukoksi lyhentävät analyysiin käytetyn työajan parista kuu-kaudesta muutamaa päivään. Jos temaattinen analyysi tehdään yksinkertaisemmin tai jätetään kokonaan väliin, lyhenee tarvittava aika muutama tuntiin.

# Lähteet

- Anderson, B., & Simpson, M. (2012). History and heritage in distance education. *Journal of Open, Flexible, and Distance Learning*, 16(2), 1-10.
- Ang, C. S. (2010). Interaction networks and patterns of guild community in massively multiplayer online games. *Social Network Analysis and Mining*, 1, 341–353.
- Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Braun, V., & Clarke, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, 12(3), 297-298.
- Daradoumis, T., Martinez-Mones, A., & Xhafa, F. (2006). A layered framework for evaluating on-line collaborative learning interactions. *Int. J. Human-Computer Studies*, 64, 622-635.
- De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46, 6-28.
- de Ruiter, P. C., Wolters, V., & Moore, J. C. (2005). *Dynamic food webs: Multispecies assemblages, ecosystem development and environmental change*. London: Academic Press.
- Dillenbourg, P., Eurelings, A., & Hakkarainen, K. (2001). Introduction. Teoksessa P. Dillenbourg, A. Eurelings, & K. Hakkarainen (toim.), *European perspectives on computer-supported collaborative learning*. Maastricht.
- Eagle, N., & Pentland, A. (2006). Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4), 255–268.
- Eckmann, J.-P., Moses, E., & Sergi, D. (2004). Entropy of dialogues creates coherent structures in e-mail traffic. *Proceedings of the National Academy of Sciences*, 101(40), 14333–14337.
- Edwards, M., Perry, B., & Janzen, K. (2011). The making of an exemplary online educator. *Distance Education*, 32(1), 101–118.
- Estrada, E. (2012). *The structure of complex networks: Theory and applications*. Oxford University Press.

- Fjermestad, J., & Hiltz, S. R. (2005). Effectiveness for students: Comparisons of “in-seat” and aln courses. Teoksessa S. R. Hiltz & R. Goldman (toim.), *Learning together online, research in asynchronous learning networks*. New Jersey: Lawrence Erlbaum Associates.
- Gunawardena, C. N., Lowe, C. A., & Andersson, T. (1997). Analysis of a global on-line debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*, 17(4), 397-431.
- Heo, H., Lim, K. Y., & Kim, Y. (2010). Exploratory study on the patterns of online interaction and knowledge co-construction in project-based learning. *Computers & Education*, 55, 1383-1392.
- Hiltz, S. R., & Goldman, R. (2005). What are asynchronous learning networks. Teoksessa S. R. Hiltz & R. Goldman (toim.), *Learning together online, research in asynchronous learning networks*. New Jersey: Lawrence Erlbaum Associates.
- Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519(3), 97-125.
- Iiskala, T., Volet, S., Lehtinen, E., & Vauras, M. (2015). Socially shared metacognitive regulation in asynchronous cscl in science: Functions, evolution and participation. *Frontline Learning Research*, 3(1), 78-111.
- Iribarren, J. L., & Moro, E. (2009). Impact of human activity patterns on the dynamics of information diffusion. *Physical Review Letters*, 103(3), 8-11.
- Isella, L., Romano, M., Barrat, A., Cattuto, C., Colizza, V., van den Broeck, W., ... Tozzi, A. E. (2011). Close encounters in a pediatric ward: Measuring face-to-face proximity and mixing patterns with wearable sensors. *PLoS ONE*, 6(2).
- Jeong, A. C. (2003). The sequential analysis of group interaction and critical thinking in online. *American Journal of Distance Education*, 17(1), 25-43.
- Järvelä, S., Malmberg, J., & Koivuniemi, M. (2016). Recognizing socially shared regulation by using the temporal sequences of online chat and logs in cscl. *Learning and Instruction*, 42, 1-11.
- Koponen, I. T. (2014). *Fysiikan käsitteet ja käsiterakenteet: Didaktisen fysiikan näkökulma*. Helsinki: Helsingin yliopisto.
- Koponen, I. T., & Nousiainen, M. (2016). Formation of reciprocal appreciation patterns in small groups: an agent-based model. *Complex Adaptive Systems Modeling*, 4(1), 24.
- Koponen, I. T., & Nousiainen, M. (2018). An agent-based model of discourse pattern formation in small groups of competing and cooperating members. *Journal of Artificial Societies and Social Simulation*, 21(2).
- Laaksonen, S.-M., & Matikainen, J. (2013). Tutkimuskohteena vuorovaikutus ja keskustelu verkossa. Teoksessa S.-M. Laaksonen, J. Matikainen, & M. Tikka (toim.), *Otteita*

- verkosta: Verkon ja sosiaalisen median tutkimusmentelmät*. Tampere: Vastapaino.
- Larson, J. R. J. (2010). *In search of synergy in small group performance*. New York: Psychology Press.
- Lawler, E., Thye, S., & Yoon, J. (2000). Emotion and Group Cohesion in Productive Exchange. *American Journal of Sociology*, 106(3), 616–657.
- Lawler, E. J. (2001). An affect theory of social exchange. *American Journal of Sociology*, 107(2), 321–352.
- Lawler, E. J., Thye, S. R.,  
Yoon, J. (2008). Social Exchange and Micro Social Order - ProQuest. *American Sociological Review*, 73(4), 519–542.
- Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction*, 13, 487–509.
- Lucas, M., Gunawardena, C., & Moreira, A. (2014). Assessing social construction of knowledge online: A critique of the interaction analysis model. *Computers in Human Behavior*, 30, 574–582.
- Lusher, D., Kremer, P., & Robins, G. (2014). *Cooperative and Competitive Structures of Trust Relations in Teams* (osa 45).
- McDonnell, M. D., Ömer Nebil Yaveroğlu, Schmerl, B. A., Iannella, N., & Ward, L. M. (2014). Motif-role-fingerprints: The building- blocks of motifs, clustering-coefficients and transitivities in directed networks. *PLoS ONE*, 9(12).
- Parke, K. (2017). Lay theories regarding computer-mediated communication in remote collaboration. *Open Praxis*, 9(1), 17-30.
- Pena-Shaffa, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, 42, 243-265.
- Przytycka, T. M., Singh, M., & Slonim, D. K. (2010). Toward the dynamic interactome: It's about time. *Briefings in Bioinformatics*, 11(1), 15–29.
- Puntambekar, S., Erkens, G., & Hmelo-Seilver, C. (2011). Introduction. Teoksessa S. Puntambekar, G. Erkens, & C. Hmelo-Seilver (toim.), *Analyzing interactions in cscs: Methods, approaches and issues*. New York: Springer.
- Savolainen, V. (2001). *Verkkoteoria*. Jyväskylä: Docendo.
- Scardamalia, M., Bereiter, C., Mclean, R. S., Swallow, J., & Woodruff, E. (1989). Computer-supported intentional learning environments. *J. Educational Computing Research*, 5(1), 51-68.
- Schellens, T., & Valcke, M. (2005). Collaborative learning in asynchronous discussion groups: What about the impact on cognitive processing? *Computers in Human Behavior* 21 (2005) 957–975, 21, 957–975.

- Schellens, T., & Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers & Education*, 46, 349-370.
- Skvoretz, J., & Fararo, T. J. (1996). Status and participation in Task Groups: A Dynamic Network Model. *The American Naturalist*, 147(4), 527-541.
- Skvoretz, J., Faust, K., & Fararo, T. J. (1996). Social structure, networks, and E-state structuralism models. *Journal of Mathematical Sociology*, 21(1-2), 57-76.
- Smedley, R. M., & Coulson, N. S. (2017). A thematic analysis of messages posted by moderators within health-related asynchronous online support forums. *Patient Education and Counseling*, 100, 1688-1693.
- Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science*, 32(11), 1492-1512.
- Tsai, C.-C., Chuang, S.-C., Liang, J.-C., & Tsai, M.-J. (2011). Self-efficacy in Internet-based Learning Environments: A Literature Review. *Educational Technology & Society*, 14(4), 222-240.
- Vaismoradi, M., Turunen, & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences*, 15, 398-405.
- Valencia, M., Martinerie, J., Dupont, S., & Chavez, M. (2008). Dynamic small-world behavior in functional brain networks unveiled by an event-related networks approach. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 77(5), 1-4.
- Vazquez, A., Rácz, B., Lukács, A., & Barabási, A. L. (2007). Impact of non-poissonian activity patterns on spreading processes. *Physical Review Letters*, 98(15), 1-4.
- Wang, Q., Woo, H. L., & Zhao, J. (2009). Investigating critical thinking and knowledge construction in an interactive learning environment. *Interactive Learning Environments*, 17(1), 95-104.
- Wasserman, S. (1994). *Social network analysis: methods and applications* (K. Faust, toim.). Cambridge: Cambridge University Press.
- Yoon, J., Thye, S. R., & Lawler, E. J. (2013). Exchange and cohesion in dyads and triads: A test of Simmel's hypothesis. *Social Science Research*, 42(6), 1457-1466.
- Zweig, K. A. (2016). *Network analysis literacy: a practical approach to the analysis of networks*. Vienna: Springer.



## Liite A. Python-skriptit

### Dialogirakenne ja roolisummat

iPython scriptit verkkokeskustelun dialogirakenteen visualisoimiseen ja roolisummien laskemiseen.

Lukee csv-tiedoston ja tekee sen pohjalta tex-tiedoston, jonka avulla latexilla saadaan piiretty asynkroninen temporaalinen verkosto. Laskee myös roolisummat ja normittaa ne sekä tallentaa niiden pohjalta tehdyt heatmap-kuvat png-tiedostoina. csv-tiedoston tulee olla muotoa:

#	Aika1	Kenelle	Aika2	Kuka	Episodi	Teema
1	1	A2	4	A1	1	1b
...	...	...	...	...	...	...

```
In [1]: ### Roolisummat, kutsu funktiota roolisummat(Adjacency_matrix). Muu  
t vain sen apuna ###  
def Rooli(Adjacency_matrix, idx): # Palauttaa yksittäisen roolin (/  
30) lukumäärän viereisyysmatriisista laskettuna  
    I = np.identity(len(Adjacency_matrix)) # yksikkömatriisi  
    U = np.ones((len(Adjacency_matrix), 1)) # yksikkövektori  
    R = Adjacency_matrix*Adjacency_matrix.T # Matrix of reciprocal  
edges  
    return{  
        1:Adjacency_matrix.T@(Adjacency_matrix-I)@U,  
        2:(Adjacency_matrix.T@Adjacency_matrix.T-R)@U,  
        3:(Adjacency_matrix.T@R-R*Adjacency_matrix)@U,  
        4:(Adjacency_matrix@Adjacency_matrix-R)@U,  
        5:Adjacency_matrix@(Adjacency_matrix.T-I)@U,  
        6:(Adjacency_matrix@R-Adjacency_matrix*R)@U,  
        7:(R@Adjacency_matrix-R*Adjacency_matrix)@U,  
        8:(R@Adjacency_matrix.T-R*Adjacency_matrix)@U,  
        9:R@(R-I)@U,  
        10:0.5*(Adjacency_matrix.T@U)*(Adjacency_matrix.T-I)@U,  
        11:((Adjacency_matrix.T@Adjacency_matrix)*Adjacency_matrix.  
T)@U,  
        12:0.5*((Adjacency_matrix.T@R)*Adjacency_matrix.T)@U,  
        13:0.5*(Adjacency_matrix@U)*(Adjacency_matrix-I)@U,  
        14:(Adjacency_matrix@Adjacency_matrix*Adjacency_matrix)@U,  
        15:0.5*((Adjacency_matrix@R)*Adjacency_matrix)@U,  
        16:(Adjacency_matrix.T@U)*(Adjacency_matrix@U)-(R@U),  
        17:((Adjacency_matrix@Adjacency_matrix.T)*Adjacency_matrix.  
T)@U,  
        18:((Adjacency_matrix@Adjacency_matrix)*Adjacency_matrix.T)  
@U,  
        19:((Adjacency_matrix@R)*Adjacency_matrix.T)@U,  
        20:(R@U)*((Adjacency_matrix.T-I)@U),  
        21:(R@Adjacency_matrix)*Adjacency_matrix.T)@U,  
        22:(Adjacency_matrix.T@Adjacency_matrix)*R)@U,  
        23:(R@R)*Adjacency_matrix.T)@U,  
        24:(R@U)*((Adjacency_matrix-I)@U),  
        25:(R@Adjacency_matrix)*Adjacency_matrix)@U,  
        26:(Adjacency_matrix@Adjacency_matrix)*R)@U,
```

```

27:((Adjacency_matrix@R)*R)@U,
28:(0.5*R@U)*((R-I)@U),
29:(R@Adjacency_matrix)*R)@U,
30:0.5*((R@R)*R)@U
}[idx]

def summat(Adjacency_matrix, roles): # Palauttaa summan useammalle
roolille
    summa=0
    for ii in roles:
        summa=summa+sum(Rooli(Adjacency_matrix, ii))
    return summa

def roolisummat(Adjacency_matrix): # Palauttaa roolisummat eli funk
tionaalisten roolien (9) kokonaislukumäärät
    summa=np.zeros(9)
    summa[0]=summat(Adjacency_matrix, [1,2,3])
    summa[1]=summat(Adjacency_matrix, [4,5,6])
    summa[2]=summat(Adjacency_matrix, [7,8,9])
    summa[3]=summat(Adjacency_matrix, [10,11,12])
    summa[4]=summat(Adjacency_matrix, [13,14,15])
    summa[5]=summat(Adjacency_matrix, [16,17,18,19])
    summa[6]=summat(Adjacency_matrix, [20,21,22,23])
    summa[7]=summat(Adjacency_matrix, [24,25,26,27])
    summa[8]=summat(Adjacency_matrix, [28,29,30])
    return summa

```

```

In [2]: ### Pikkufunctiot ###
def uniikit_nimimerkit(df): # Palauttaa DataFrameesta kaikki yksilöl
liset nimimerkit aakkosjärjestyksessä
    return np.sort(np.unique(np.concatenate((df.Kuka.unique(), df.K
enelle.unique()))))

def nimi_koodiksi(nimi): # Palauttaa nimerkkiä koodin. HUOM. Tämä o
n aineistokohtainen
    nimimerkit=('ES', 'HL', 'JS', 'MP', 'HB', 'JK', 'ME', 'MR', 'AR
', 'IR', 'PT', 'TS', 'JT', 'MA', 'OK', 'SS')
    koodit=("A1", "A2", "A3", "A4", "B1", "B2", "B3", "B4", "C1", "
C2", "C3", "C4", "D1", "D2", "D3", "D4")
    return koodit[nimimerkit.index(nimi)]

def episodi(tmp): # Episodia vastaava väri, käytetään ATV tikz-kuvi
ssa
    colors = ("black", "blue", "red", "green")
    return colors[tmp]

def nimi_idx(tmp, nimet): # Palauttaa
    return int(np.where(nimet==tmp)[0])

def df_pois(df): # Palauttaa DataFramen, jossa on vain yksi rivi vi
estiä kohden
    pois=[]
    for ii in range(1, len(df.index)):
        if df.get_value(ii-1,"#") == df.get_value(ii,"#"):
            pois.append(ii)
    df=df.drop(pois)
    df.index = range(len(df.index))

```

```
    return df

def df_vain_viestit(df): # Palauttaa DataFramen, jossa ei viestejä
    ilman vastauksia tai vastauksia itselle, teemat poistettu
    df_viestit=pd.DataFrame(columns=("Aika1", "Kenelle", "Aika2", "
    Kuka", "Episodi"))
    for ii in df.index:
        kenelle=df.get_value(ii,"Kenelle")
        kuka=df.get_value(ii,"Kuka")
        if kenelle != kuka:
            df_viestit.loc[ii]=df.loc[ii]
    df_viestit.index = range(len(df_viestit.index))
    return df_viestit

def df_episodit(df): # Palauttaa kolme DataFramea, joissa jokaisessa
    vain ko. episodin viestit.
    df1=df.loc[df["Episodi"] == 1]
    df1.index = range(len(df1.index))
    df2=df.loc[df["Episodi"] == 2]
    df2.index = range(len(df2.index))
    df3=df.loc[df["Episodi"] == 3]
    df3.index = range(len(df3.index))
    return (df1, df2, df3)

def aika_askeleen_pituus(tmp):
    if tmp < 16:
        return 1
    else:
        return int(np.round(tmp/10))
```

```

In [3]: ### Asynkroninen Temporaalinen verkosto, ATV, tänne df_pois ###
def ATV(tiedosto, df, nimimerkit):
    nimi_tex="gradu_atv_tikz/atv_v2/" + tiedosto + "_tikz.tex"
    rivit=df.shape[0] # Rivien lukumäärä. Tikzin matriisissa yksi
tai kaksi riviä enemmän
    df=df.sort_values(by=['Episodi']) # Sorttaus episodeittain
    df.index = range(len(df.index)) # indexit 0,1,2...
    nimet=""
    for ii in nimimerkit: # Korvataan nimikirjaimet koodeilla
        nimet=nimet+ " & " + nimi_koodiksi(ii) #nimimerkit korvatan
koodeilla A1...D4
    nimet="t"+nimet+" \\ \n"
    rivi=" & \\null & \\null & \\null & \\null \\ \n"
    f = open(nimi_tex,'w') # Avataan tekstitiedosto
    f.write(nimet) # rivien otsikot
    for ii in range(1,df["Aika2"].max()+1): # Null-matriisi
        rivi_ii=str(ii)+rivi
        f.write(rivi_ii)
    f.write(" \\null & \\null & \\null & \\null & \\null \\ \n")
    f.write("}; \n")
    # Pallot ja nuolet
    rivi="\draw[my_circle](m-"
    rivi2=")circle; \n"
    rivi3="\draw [arrow, color="
    rivi4 ="] (m-"
    rivi5=") -- (m-"
    rivi6="); \n"
    for ii in range(0,rivit):
        aika1=df.get_value(ii,"Aika1")+1
        kenelle=df.get_value(ii,"Kenelle")
        kenelle=nimimerkit.tolist().index(kenelle)+2
        rivi_ii_1=rivi+str(aika1)+"-"+str(kenelle)+rivi2
        f.write(rivi_ii_1)
        aika2=df.get_value(ii,"Aika2")+1
        kuka=df.get_value(ii,"Kuka")
        kuka=nimimerkit.tolist().index(kuka)+2
        rivi_ii_2=rivi+str(aika2)+"-"+str(kuka)+rivi2
        # Jos eka ja toka aika ja hlö tagit erit, lisätään toinenki
n pallo ja nuoli niiden välille
        if rivi_ii_1 != rivi_ii_2:
            f.write(rivi_ii_2)
            rivi_ii_3=rivi3+episodi(df.get_value(ii,"Episodi"))+rivi4+str(aika1)+"-"+str(kenelle)+rivi5+str(aika2)+"-"+str(kuka)+rivi6
            f.write(rivi_ii_3)
        # Pystyviivat
        f.write("\n \draw [dotted] (m-1-2) -- (m-"+str(df["Aika2"].max(
)+2)+"-2); \n" \
        "\draw [dotted] (m-1-3) -- (m-"+str(df["Aika2"].max()+2)+"-3); \n" \
        "\draw [dotted] (m-1-4) -- (m-"+str(df["Aika2"].max()+2)+"-4); \n" \
        "\draw [dotted] (m-1-5) -- (m-"+str(df["Aika2"].max()+2)+"-5); \n")
    f.close()

```

```
In [4]: ### Matriisisummat, eli roolisummien aikakehitys, käytä tätä episo
deittain s.e. tiedostossa on episodi###
def Matriisisummat(tiedosto, df, nimimerkit, aika_askel):
    df=df.sort_values(by=['Aika2']) # Järjestetään vastauksien per
    usteella aikajärsstyukseen ja indexit oikein
    df.index = range(1,len(df.index)+1)
    A = np.zeros((4,4)) # viereisyysmatriisin pohjustus
    columns=np.array(range(1,10)) # Roolisummien dataframet
    df_summat=pd.DataFrame(columns=columns)
    df_summat_trans=pd.DataFrame(columns=columns)
    for ii in df.index: # Lasketaan viereisyysmatriisi ja roolisum
    mat
        kenelle=df.get_value(ii,"Kenelle")
        kuka=df.get_value(ii,"Kuka")
        A[nimi_idx(kuka, nimimerkit),nimi_idx(kenelle, nimimerkit)]
    =1 # Merkitään viereisyysmatriisiin 1, sos. suhde
        if ii % aika_askel == 0 and ii != 0:
            df_summat.loc[ii]=roolisummat(A)
            df_summat_trans.loc[ii]=roolisummat(np.transpose(A)) #
    Transpoosilla tiedonsuunta
        if len(df.index) % aika_askel != 0: # Lisätään df_summat viimei
    nen rivi, jos ei täsmää aika-askeleen kanssa
            df_summat.loc[len(df.index)]=roolisummat(A)
            df_summat_trans.loc[len(df.index)]=roolisummat(np.transpose
    (A))
    return A, df_summat, df_summat_trans
```

```
In [5]: ### Roolien plottaus png:nä ###
def matriisi_plot(tiedosto, df_summat):
    fig=plt.figure(num=None, figsize=(5, 5), dpi=240, facecolor='w'
    , edgecolor='k')#figsize=(5,10))
    ax = plt.subplot()
    #max_roolit=np.max(np.max(df_summat))
    vmax=100 #max_roolit
    #cmap = plt.get_cmap('gray_r')#, vmax)
    cmap = matplotlib.colors.LinearSegmentedColormap.from_list("",
    ["white", "lightsteelblue", "black"])
    cax=ax.matshow(df_summat, cmap=cmap, vmin=0, vmax=vmax) # cmap=
    'gray_r' käänteinen colormappi
    #plt.ylabel('Aika-askel', fontsize=12) # Ei turhaan kaikkiin ak
    seleiden nimiä
    #plt.xlabel('Roolit', fontsize=12)
    #plt.title("Roolien aikakehitys")
    xaxis=range(1,10)
    yaxis=df_summat.index
    plt.xticks(range(len(xaxis)), xaxis, size=18)
    plt.yticks(range(len(yaxis)), yaxis, size=18)
    #fig.colorbar(cax)
    nimi_png="gradu_heatmap_png/png_v2/" + tiedosto + "_heatmap.png
    "
    plt.savefig(nimi_png, bbox_inches='tight')
    plt.close()
```

```
In [6]: ### Roolien plottaus normitettuna ja png:nä ###
def matriisi_plot_norm(tiedosto, df_summat):
    norm=np.array([72,72,72,48,48,96,96,96,48]) # Nämä saatu täyden
    verkon roolisummista.
    fig=plt.figure(num=None, figsize=(5, 5), dpi=240, facecolor='w'
, edgecolor='k')#figsize=(5,10))
    ax = plt.subplot()
    #max_roolit=np.max(np.max(df_summat))
    vmax=1 #max_roolit
    #cmap = plt.get_cmap('gray_r')
    cmap = matplotlib.colors.LinearSegmentedColormap.from_list("",
["white", "lightsteelblue", "black"])
    cax=ax.matshow(df_summat/norm, cmap=cmap, vmin=0, vmax=vmax) #
cmap='gray_r' käänteinen colormappi
    #plt.ylabel('Aika-askel', fontsize=12) # Ei turhaan kaikkiin ak
seleiden nimiä
    #plt.xlabel('Roolit', fontsize=12)
    #plt.title("Roolien aikakehitys")
    xaxis=range(1,10)
    yaxis=df_summat.index
    plt.xticks(range(len(xaxis)), xaxis, size=18)
    plt.yticks(range(len(yaxis)), yaxis, size=18)
    #fig.colorbar(cax)
    nimi_png="gradu_heatmap_png/png_v2/" + tiedosto + "_normitettu_
heatmap.png"
    plt.savefig(nimi_png, bbox_inches='tight')
    plt.close()
```

```
In [7]: ### Tällä voi tehdä erillisen colorbarin ###
def eril_colorbar():
    a = np.array([[0,1]])
    plt.figure()
    cmap = matplotlib.colors.LinearSegmentedColormap.from_list("",
["white", "lightsteelblue", "black"])
    img = plt.imshow(a, cmap=cmap)
    plt.gca().set_visible(False)
    cbar=plt.colorbar(orientation="horizontal")
    cbar.ax.set_aspect('auto') #Ilman tätä twiny() sotkee kuvasuhte
et
    # Toinen y-akseli ja sen rajat
    ax2 = cbar.ax.twiny()
    ax2.set_xlim([0,100])
    plt.savefig("colorbar.png",bbox_inches='tight') #Tallenna
```

```

In [8]: ### Varsinainen Looppi, jossa tehdään kaikki yllä oleva kaikille .c
sv-tiedotoille ###
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.colors
from matplotlib import rc
# Muut fontit: https://matplotlib.org/users/usetex.html
rc('font',**{'family':'serif','serif':['Computer Modern Roman']})
rc('text', usetex=True)

print("looppi aloitettu")
for ii in range(1,5):
    for jj in range(1,5):
        nimi="h"+str(ii)+"r"+str(jj)+"_kopio" # Tiedoston nimi
        nimi_csv=nimi + ".csv"
        df = pd.read_csv(nimi_csv) # Luetaan tiedosto dataframeksi
        df=df_pois(df) # Ylimääräiset rivit pois
        nimimerkit=uniikit_nimimerkit(df) # Uniikit nimimerkit aako
sjärjestyksessä
        df_viestit=df_vain_viestit(df) # DF jossa ei vastauksia its
elle tai viestejä ilman vastauksia
        dff=df_episodit(df_viestit) # Jaetan episodeittain
        ATV(nimi, df, nimimerkit) # .tex tiedosto Asynkronisen temp
oraaliverkoston piirtämiseen tikzillä
        for kk in range(0,3): # Lasketaan ja plotataan roolisummat
episodeittain
            if len(dff[kk]) > 5: #millä viestien määrällä lasketaan
roolisummat molempiin suuntiin
                nimi_ep=nimi+"_ep_"+str(kk+1) # Nimet .png tiedosto
ille

                nimi_trans=nimi+"_ep_"+str(kk+1) + "_trans_"
                A, df_summat, df_summat_trans = Matriisisummat(nimi
, dff[kk], nimimerkit, aika_askeleen_pituus(len(dff[kk])))
                matriisi_plot(nimi_ep, df_summat)
                matriisi_plot(nimi_trans, df_summat_trans)
                matriisi_plot_norm(nimi_ep, df_summat)
                matriisi_plot_norm(nimi_trans, df_summat_trans)
                print("h"+str(ii)+"r"+str(jj)+ " episodi " + str(kk
+1) + " valmis")
eril_colorbar()
print("looppi valmis")

```

```
looppi aloitettu
h1r1 episodi 1 valmis
h1r1 episodi 2 valmis
h1r2 episodi 1 valmis
h1r3 episodi 1 valmis
h1r4 episodi 1 valmis
h2r1 episodi 1 valmis
h2r1 episodi 3 valmis
h2r2 episodi 1 valmis
h2r2 episodi 2 valmis
h2r3 episodi 1 valmis
h2r3 episodi 2 valmis
h2r4 episodi 1 valmis
h2r4 episodi 2 valmis
h3r1 episodi 1 valmis
h3r1 episodi 2 valmis
h3r1 episodi 3 valmis
h3r2 episodi 1 valmis
h3r2 episodi 2 valmis
h3r3 episodi 1 valmis
h3r4 episodi 1 valmis
h3r4 episodi 2 valmis
h4r1 episodi 1 valmis
h4r2 episodi 1 valmis
h4r2 episodi 2 valmis
h4r3 episodi 2 valmis
h4r4 episodi 1 valmis
h4r4 episodi 2 valmis
looppi valmis
```

```
In [9]: # Tällä solut täysikokoiseksi
from IPython.core.display import display, HTML
display(HTML("<style>.container { width:100% !important; }</style>"
))
```



## Pylväsdiagrammit

Scripti joka laskee csv-tiedostoista episodikohtaiset teemat ja tekee niistä tex-tiedoston, jota voidaan käyttää Latexissa pylväsdiagrammien piirtämiseen. Laskee ja tulostaa myös yhteenlasketujen teemojen lukumäärät.

In [1]:

```
def counts_tex(filename, counts):
    filename=filename+".tex"
    f = open(filename, 'w')
    f.write("\addplot+[bar_a!20!black,fill=bar_a] coordinates { (" +str(counts
["0"])+",0) (" +str(counts["1a"])+",1) (" +str(counts["2a"])+",2) (" +str(counts[
"3a"])+",3) (" +str(counts["4a"])+",4) (" +str(counts["5"])+",5) (" +str(counts["
6"])+",6) (" +str(counts["7"])+",7)}; "+"\\n")
    f.write("\addplot+[bar_b!20!black,fill=bar_b] coordinates { (0,0) (" +str(
counts["1b"])+",1) (" +str(counts["2b"])+",2) (" +str(counts["3b"])+",3) (" +str(
counts["4b"])+",4) (0,5) (0,6) (0,7)}; "+"\\n")
    f.write("\addplot+[bar_c!20!black,fill=bar_c] coordinates { (0,0) (" +str(
counts["1c"])+",1) (" +str(counts["2c"])+",2) (" +str(counts["3c"])+",3) (" +str(
counts["4c"])+",4) (0,5) (0,6) (0,7)}; "+"\\n")
    f.write("\addplot+[bar_d!20!black,fill=bar_d] coordinates { (0,0) (" +str(
counts["1d"])+",1) (" +str(counts["2d"])+",2) (0,3) (" +str(counts["4d"])+",4) (
0,5) (0,6) (0,7)}; "+"\\n")
    f.write("\addplot+[bar_e!20!black,fill=bar_e] coordinates { (0,0) (" +str(
counts["1e"])+",1) (" +str(counts["2e"])+",2) (0,3) (0,4) (0,5) (0,6) (0,7)}; "+"
\\n")
    f.write("\addplot+[bar_f!20!black,fill=bar_f] coordinates { (0,0) (0,1) (
"+str(counts["2f"])+",2) (0,3) (0,4) (0,0) (0,6) (0,7)}; "+"\\n")
    f.close()

def df_episodit(df): # Palauttaa kolme DataFramea, joissa jokaisessa vain ko. e
pisodin viestit.
    df1=df.loc[df["Episodi"] == 1]
    df1.index = range(len(df1.index))
    df2=df.loc[df["Episodi"] == 2]
    df2.index = range(len(df2.index))
    df3=df.loc[df["Episodi"] == 3]
    df3.index = range(len(df3.index))
    return (df1, df2, df3)

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from collections import Counter

kaikkiteemat=[]
for ii in range(1,5): # Luetaan kaikista csv-tiedostoista DataFrame
    for jj in range(1,5):
        nimi="h"+str(ii)+"r"+str(jj)+" kopio" # Tiedoston nimi
        nimi_csv=nimi + ".csv"
        df = pd.read_csv(nimi_csv)
        kaikkiteemat.extend(df["Teema"].tolist()) # Lisätään kaikkiin teemoihin
```

```

n
    dff=df_episodit(df)
    for kk in range(1,4):
        nimi_ep="bar_charts/"+nimi+" ep"+str(kk)
        tmp=[]
        df_ep=dff[kk-1]
        tmp.extend(df_ep["Teema"].tolist()) # Jatketaan listaa listalla te
emoista
        if len(tmp)!=0:
            counts = Counter(sorted(tmp)) # Teemojen lukumäärät
            filename="bar_charts/bar_r"+str(jj)+"h"+str(ii)+"ep"+str(kk)
            counts_tex(filename, counts)

kaikkiteemat = Counter(sorted(kaikkiteemat)) # Kaikkien teemojen lukumäärä
kaikkiteemat

```

Out[1]:

```

Counter({'0': 56,
        '1a': 176,
        '1b': 124,
        '1c': 31,
        '1d': 90,
        '1e': 83,
        '2a': 18,
        '2b': 7,
        '2c': 9,
        '2d': 5,
        '2e': 41,
        '2f': 95,
        '3a': 6,
        '3b': 19,
        '3c': 7,
        '4a': 20,
        '4b': 195,
        '4c': 34,
        '4d': 34,
        '5': 6,
        '6': 44,
        '7': 54})

```

In [2]:

```

# Tällä solut täysikokoiseksi
from IPython.core.display import display, HTML
display(HTML("<style>.container { width:100% !important; }</style>"))

```

## Scripti moodle foorumin purkamiseen xml-tiedostosta

Moodle-alustan foorumikeskustelut saa xml-tiedostona tekemällä kurssista varmuuskopion. Huom, varmuuskopion tiedostopääte pitää muuttaa .zip:ksi, jotta sen voi purkaa. Kansiota löytyy tiedosto activities/forum\_n/forum.xml jossa n on foorumin/modulen id-numero.

Tämä scripti lukee tuon xml-tiedoston ja tuottaa kaksi csv-tiedostoa, joiden nimissä käytetään samaa id-numeroa kuin kansiossakin. Ensimmäisessä csv-tiedostossa on vuorovaikutukset eli kuka on vastannut kenellekin ja millä ajanhetkellä. Nämä ovat aika järjetyksessä lähetetyn viestin ajan mukaan (T1). Toisessa csv-tiedostossa on lähetetyt viestit käyttäjäkohtaisissa sarakkeissa aikajärjestyksessä.

Scriptissä vaihdetaan xml-tiedostossa olevat unix-ajat tunti:minuutti -muotoon. unix-aika sisältyy myös päivämäärän ja kellonajat sekunnin tarkkuudella ja niitä onkin käytetty tietojen järjestämiseen ts. jos tiedostossa on peräkkäin kaksi samaa aikaa minuutin tarkkuudella on jälkimmäinen viesti lähetetty jotain sekunteja myöhemmin. Myös päivämäärän vaihtuminen havaitaan kellon mennessä ympäri.

Tiedostoissa ei ole käyttäjiä nimetty, vaan he esiintyvät tunnistenumeroiden kautta. Nämä tunnistenumerot saa varmuuskopion tiedostosta users.xml ja toinen scripti tekee näistä yksinkertaisen csv:n.

In [ ]:

```
#####
#####
#### Scripti moodle foorumin purkamiseen xml-tiedostosta. Miikka Turkkila, sy
ksy 2017 ####
#####
#####

### Funktiot ###
def parentti(postid, arg2): # Etsii postid:n avulla parentin tai parentin post
aus ajan. Jos ei parenttia, palauttaa alkuperäisen viestin useridn tai postau
sajan
    Post=soup.find("post", id=postid) # Mikä viesti
    Parentid=int(Post.parentid.text)
    if Parentid==0:
        if arg2=="time":
            return int(Post.created.text)
        elif arg2=="who":
            return int(Post.userid.text)
    else:
        Parent=soup.find("post", id=Parentid)
        if arg2=="time":
            return int(Parent.created.text)
        elif arg2=="who":
            return int(Parent.userid.text)

def userit(soup): # löytää soupista kaikki uniikit user idt a palauttaa ne int
teinä järjestettynä
    seen = []
```

```

    for tmp in soup.find_all("post"):

        if int(tmp.userid.text) not in seen:
            seen.append(int(tmp.userid.text))
        seen.sort()
    return seen

def kaikki_ajat(soup): # etsii kaikki postaus ajat soupista
    kaikkiajat=[]
    for tmp in soup.find_all("created"):
        kaikkiajat.append(int(tmp.text))
    return kaikkiajat

#### Funktiot loppuu #####
#####

from datetime import datetime
from bs4 import BeautifulSoup
from pandas import DataFrame, read_csv

with open("forum kopio.xml") as data: # Avataan .xml-tiedosto soup-objektiksi
    soup = BeautifulSoup(data, "xml")

# Vaihtaa kaikki parent-tagit parentid-tageiksi, koska .parent on metodi bs4:ssä
while True:
    tag = soup.find("parent")
    if not tag:
        break
    tag.name = "parentid"

# Poistaa turhat html-tagit viesteistä, näitä saattaa tulla lisää
for tmp in soup.find_all("message"):
    viesti=tmp.text
    viesti=viesti.replace("<p>", "")
    viesti=viesti.replace("</p>", "")
    viesti=viesti.replace("<br>", "\n") # Linebreak
    viesti=viesti.replace("&nbsp;", "") # ylimääräinen väli
    viesti=viesti.replace("&lt;", "<")
    #viesti=viesti.replace("&", "&")
    tmp.string=viesti

# Perustietoja
post_n=len(soup.find_all("post")) #postauksien lukumäärä
# discussion_n=len(soup.find_all("discussion")) #keskusteluiden lukumäärä
moduleid=soup.activity["moduleid"] # modulen/forumin id csv:n tiedostonimeen

# Dataframien alustus. Kaksi dataframea, yksi vuorovaikutuksille ja toinen sisällöille
indexit=range(1,post_n+1)
columns = ['T1','Kenelle','T2','Kuka','Episodi','Aihe','Viesti'] #Aihe eli keskustelun otsikko. Episodin ideana
df_vv = DataFrame(index=indexit, columns=columns)

opp=userit(soup)
ajat=kaikki_ajat(soup)
df_sis = DataFrame(index=indexit, columns=opp)
df_sis.insert(0, "T", ajat)
# Dataframet alustettu

```

```

# Samassa loopissa täydennetään molemmat dataframet
# Looppi joka käy kaikki discussionit, discussion id = episodi, voi korvata 1,
2,3... myöhemmin
index=0
for discussion in soup.find_all("discussion"):
    discussion_id=discussion["id"]
    post=discussion.post # discussionin eka postaus
    #Looppi joka käy kaikki discussionin postaukset läpi ja kirjoittaa halutut
    tiedot molempiin dataframeihin
    for ii in range(0,len(discussion.find_all("post"))): #Miten range kohillee
n? Vähän purkka-jesari-viritys
        index+=1
        # etsitään soupista tarvittavat tiedot
        PostID=post["id"]
        UserID=post.userid.text
        Created=int(post.created.text)
        ParentID=post.parentid.text
        Message=post.message.text
        Aihe=post.subject.text
        Kenelle=parentti(int(PostID), "who")
        t1=int(parentti(int(PostID), "time"))
        # Viestit sisällöt dataframeen. Tässä int(UserID), koska df_sis opp-co
lumnit intteinä
        df_sis.loc[df_sis.loc[df_sis['T'] == Created].index[0], int(UserID)]=M
essage
        # Arvot vuorovaikutus dataframeen
        df_vv.loc[index, 'T1']=t1 # T1
        df_vv.loc[index, 'Kenelle']=Kenelle # Kenelle
        df_vv.loc[index, 'T2']=Created # T2
        df_vv.loc[index, 'Kuka']=UserID # Kuka
        df_vv.loc[index, 'Episodi']=discussion_id # Episodi
        df_vv.loc[index, 'Viesti']=Message # Viesti
        df_vv.loc[index, 'Viesti']=Message # Viesti
        df_vv.loc[index, 'Aihe']=Aihe # Aihe
        # Löytää seuraavan postauksen
        post=post.find_next("post")

# Sorttaa dataframet ajan mukaan, ihan varmuuden vuoksi ja asettaa indexit uud
estaan 1,2,3...
df_vv.sort_values(["T1", "T2"], inplace=True)
#df_vv.reset_index(drop=True, inplace=True)
df_vv.index = indexit
df_sis.sort_values("T", inplace=True)
#df_sis.reset_index(drop=True, inplace=True)
df_sis.index = indexit

# Unix ajat kellonajoiksi
for ii in indexit:
    df_vv.loc[ii, "T1"]=datetime.fromtimestamp(df_vv.loc[ii, "T1"]).strftime('
%H:%M') # '%d.%m.%y %H:%M:%S'
    df_vv.loc[ii, "T2"]=datetime.fromtimestamp(df_vv.loc[ii, "T2"]).strftime('
%H:%M')
    df_sis.loc[ii, "T"]=datetime.fromtimestamp(df_sis.loc[ii, "T"]).strftime('
%H:%M')

# Dataframet csv-tiedostoiksi. Tiedostojen nimiin forum id:t

```

```

vv_csv="vuorovaikutukset_forum_"+moduleid+".csv"
df_vv.to_csv(vv_csv,index=True,header=True)
print(vv_csv, "-tiedosto luotu")
sis_csv="sisällöt_forum_"+moduleid+".csv"
df_sis.to_csv(sis_csv,index=True,header=True)
print(sis_csv, "-tiedosto luotu")

```

## Käyttäjät

Tämä scripti tuottaa csv-tiedoston käyttäjien nimistä, id-tunnisteista ja rooleista. Tämä tarvitsee users.xml ja course/roles.xml -tiedostot toimiakseen. Toisesta löytyy nimet yhdistettynä id:ihin ja toisesta roolit. Roolit ovat tiedostoissa vain numeroina, mutta tiedostosta roles.xml löytyy kuvaukset. Hölmösti kahdella eri tiedostolla sama nimi, mutta toisessa rooli kuvaukset ja löytyy suoraan varmuuskopiokansioista ja varsinaiset roolit course-alikansiosta. Lisätään non-editing teacher -rooli, jos tarpeellista. Jos henkilöä ei löydy course/roles.xml -tiedostosta, on roolina "Ei roolia". Syytä tähän epäohjonmukaisuuteen en keksi, mutta ehkä henkilö poistuu kurssilta.

In [ ]:

```

#####
###
#### Scripti, joka yhdistää userid:t käyttäjiin. Miikka Turkkila, syksy 2017 #
###
#### Tarvitsee users.xml ja roles.xml -tiedosto. Löytyy varmuuskopiosta      #
###
#####
###
def userit(soup): # löytää soupista kaikki uniikit user idt a palauttaa ne int
teinä järjestettynä
    seen = []
    for tmp in soup.find_all("user"):
        tmp2=tmp["id"]
        if int(tmp2) not in seen:
            seen.append(int(tmp2))
    seen.sort()
    return seen

def rooli(soup, tmp): # Etsii roolin.
    tmp2=soup.find("userid", string=str(tmp))
    if tmp2!=None:
        rooli=int(tmp2.find_previous("roleid").text)
        if rooli==3:
            return "Opettaja"
        elif rooli==5:
            return "Opiskelija"
    else:
        return "Ei roolia"

from bs4 import BeautifulSoup
from pandas import DataFrame, read_csv

with open("users kopio.xml") as data: # Avataan .xml-tiedosto soup-objektiksi
    soup id = BeautifulSoup(data, "xml")

```

```
with open("roles kopio.xml") as data: # Avataan .xml-tiedosto soup-objektiksi
    soup_roles = BeautifulSoup(data, "xml")

users=userit(soup_id)
user_n=len(users)
# Dataframe
indexit=range(1,user_n+1)
columns = ['sukunimi','etunimi', 'rooli']
df_users = DataFrame(index=indexit, columns=columns)
df_users.insert(0, "id", users)

for tmp in users:
    idx=df_users.loc[df_users['id'] == tmp].index[0]
    df_users.loc[idx, "sukunimi"]=soup_id.find("user", id=tmp).lastname.text
    df_users.loc[idx, "etunimi"]=soup_id.find("user", id=tmp).firstname.text
    df_users.loc[idx, "rooli"]=rooli(soup_roles, tmp)

users_csv="osallistujat.csv" #Tähän ei kivasti saa forumin id:tä
df_users.to_csv(users_csv,index=True,header=True)
print(users_csv, "-tiedosto luotu")
```